

Building Up Social Capital in a Changing World

Fernando Vega-Redondo

Departamento de Fundamentos del Análisis Económico and
Instituto Valenciano de Investigaciones Económicas
Universidad de Alicante, 03071 Alicante, Spain

and

Departament de Economia i Empresa
Universitat Pompeu Fabra, 08005 Barcelona, Spain

Preliminary version: September 2002

Abstract

This paper models the dynamic process through which a large society may succeed in building up its “social capital” by establishing a stable and dense pattern of interaction among its members. In the model, agents interact according to a collection of (idiosyncratic) infinitely repeated Prisoner’s Dilemma played on the existing social network. This network not only specifies the playing partners but, crucially, also determines how relevant strategic information diffuses or new cooperation opportunities are found. Over time, the underlying payoffs randomly change, i.e. display some “volatility”. In response to it, agents react by creating new links and removing others. This combines into a complex but ergodic dynamic process, whose analysis is undertaken in different ways. First, we rely on its ergodicity to “compute” numerically its long-run regularities. Second, we use mean-field approximations to derive analytical results. Both routes are found in accord and also complementary.

The long-run dynamics of the process sharply depends on environmental volatility, displaying the following features: (a) Only if volatility is not too high can the society sustain a dense social network and thus attain a large average payoff. (b) The social architecture endogenously responds to increased volatility by becoming more cohesive. (c) Network-based strategic effects are an essential buffer that preclude the abrupt collapse of the social network in the face of growing volatility. These conclusions are largely in tune with the points stressed in the social-capital literature.

1 Introduction

Within any socio-economic environment, the network of social interaction is the backbone that underlies its overall operation. The role of the social network is two-fold. On the one hand, of course, it determines how agents come into contact to carry out their economic activities. But, complementary to this, the social network also maps how the relevant information required to conduct those activities flows among the agents. To shed light, therefore, on the interplay between behavior and information that is at the core of so many social phenomena, it is important to have a proper understanding of the main forces that underlie the network of social interaction.

In principle, this network should be conceived as an endogenous outcome of agents' decisions, much in the same way as any other dimension of choice, e.g. whether to play a certain game or undertake some investment. Networking decisions, however, are particularly interesting in that they display the following features:

- (i) A new social link is typically initiated and supported by a small collection of players – say, bilaterally – but it may induce important externalities on the general social system. For example, it may open new channels of communication that could prove crucial for implementing wide social gains.
- (ii) The establishment and maintenance of links is costly (e.g. time consuming) and therefore “investment” in them should respond to the usual economic considerations, i.e. opportunity costs.
- (iii) The accumulation of links is bound to be only gradual, since the adjustment costs entailed by a very fast rate of link formation are typically prohibitive.
- (iv) Preexisting links depreciate in that, for example, the value of the activity and interaction they support may become obsolete or/and not so rewarding (relative to other options) as it used to be.

The features listed above are standard in many problems of investment and suggest conceiving the process of network formation as the accumulation of a special kind of capital – social capital. The term “social capital” has been used in recent times with a variety of different meanings, some of them perhaps too vague or devoid of operational content.¹ Here, I focus on one of the most widely agreed incarnations of this concept. I identify the stock of social capital enjoyed by a certain community with the density and stability of its social network. This, of course, is motivated by the implicit assumption that some dense and stable

¹Even though some have pointed to an earlier origin, it was arguably James Coleman (see Coleman (1988)) who brought the notion of social capital to prominence in the sociological, and then economic, literature. For a good and recent overview on the use and possible misuse of this concept, see Woolcock (2000), whose discussion mostly focuses on development issues.

interaction has positive welfare implications, and should typically be correlated with high overall payoffs. (Admittedly, this assumption may not be suitable for some applications, as stressed, for example, by Durlauf (1999).)

To address the aforementioned issues in a simple and paradigmatic context, I propose a model where players are involved in a collection of pairwise Infinitely Repeated Prisoner's Dilemmas (IRPD). Every pair of agents which happens to be linked in the social network play an *idiosyncratic* version of this game (i.e. cooperation and relative opportunistic gains typically differ across pairs). These games are played independently, in the sense that the choices made in each of them (cooperation or defection) are adopted independently at each stage by the players involved. The different games, however, are not *strategically* independent since the behavior of a player in one of the games she plays can be made dependent on what has previously happened in other games.² Such information on past behavior, however, is not assumed to diffuse instantaneously. Rather, it is supposed to "travel" gradually (one step/link at a time) along the network. Of course, only when it arrives to any particular player can the latter's choice be affected by it – say, triggering a punishment to a then-revealed defector.

In this context, it is apparent that the range of incentive-compatible behavior that can be supported in the infinitely repeated population game must be crucially dependent on the architecture of the underlying network. And reciprocally, of course, the particular network that should prevail – more specifically, which links will be formed and which removed – also has to depend on the payoffs that can be earned in an incentive-compatible fashion. To formalize these considerations, I define the notion of *Pairwise-Stable Network* (PSN), that combines standard ideas from the literature on repeated games with the concept of pairwise stability found in the matching and network-formation literatures. Informally, a PSN is a network in which each of its extant links supports bilateral cooperation when every player uses optimal trigger strategies in all of her (repeated) interactions.

In line with the dynamic approach of the paper, a PSN may be conceived as the rest point of an adjustment process resulting from the accumulation of revision opportunities enjoyed by different pairs of connected players. Thus, the first step undertaken in the paper is to *characterize* such networks. I find that rather fine details of the architecture of the network, in conjunction with players' strategic incentives in the repeated games, are important to understand pairwise stability. For example, a key consideration that generally underpins the stability of a link (i.e. the incentives of both players to cooperate and thus keep their connection) is the existence of other "valuable" neighbors who are not too distant apart and could punish either deviation without much delay. In other

²A related approach can be found in Bernheim and Whinston (1990). These authors explore whether multimarket interaction enhances the collusion possibilities of an oligopoly. The key difference with the present approach is that, in their case, the same set of firms is assumed to interact in all markets simultaneously and, consequently, the information on the past behavior of other oligopolists may be brought to bear immediately on all markets. That is, in their case, the strategically relevant information spreads instantaneously throughout all interactions.

words, some measure of network cohesiveness (or “cliquishness”) turns out to be important in supporting network stability.

As explained, however, the primary objective of the paper is not static but dynamic. Our aim, therefore, is to shed light on how, over time, the strategic incentives to cooperate interplay with the considerations underlying the formation of the social network. To this end, an adjustment (evolutionary) process is postulated that is taken to proceed on a “slow” time scale, relative to the rate at which the stage game is repeated. The main features of this process can be summarized as follows.

First, as a basic step in the adjustment, and prior to any actual play, every pair of connected players is given the opportunity to check the stability of their respective links. Doing so, only the links which are able to support bilateral cooperation are maintained. If this mechanism were to operate by itself, it is quite clear that it would eventually lead to a pairwise-stable network (possibly empty) that is a certain subset of the original one. Thus, to enrich the dynamics, we add, in every period, two sources of “change and innovation”:

- (1) *Update of payoff conditions*: the payoffs of existing links are changed with some (small) independent probability, say $\varepsilon > 0$. This probability – a key parameter of the model – is interpreted as a stylized measure of environmental volatility.
- (2) *Search and link creation*: Each player can support a maximum pre-specified number of total links. Within this limitation, however, every player receives (with some independent probability) the opportunity of forming a fresh new link with some new player she “knows” – i.e. with someone to whom she is connected through a network path.

By making search and link creation a network-bound task, the process induced by (1)-(2) displays acute path dependencies that render initial conditions overly powerful in the determination of long-run behavior. (For example, an initially empty social network would remain so forever, since no individual ever “knows” any other one with whom to connect.) This seems too-rigid a formulation and suggests extending (2) as follows:

- (2⁺) *Enhanced search and link creation*: In addition to (2), with some (small) independent probability, every player enjoys the possibility of forming a new link with a previously “unknown” player – i.e. with someone in another component of the social network.

The stochastic process where the removal of pairwise-unstable links is complemented by payoff update (1) and enhanced search (2⁺) will be proven to be ergodic. Its long-run behavior, therefore, is univocally summarized, *independently* of initial conditions, by its unique invariant distribution. This invariant distribution will be characterized in some particular scenarios – for example, when the range of admissible payoffs is low or when the payoff environment is stable (i.e. $\varepsilon = 0$). These parameter configurations represent an useful benchmark of comparison for the general case. For less extreme scenarios, however,

an analytical characterization of the long-run behavior seems unfeasible. But then, the fact that the process is ergodic affords an indirect way of “computing” the long-run invariant magnitudes of any variable of interest. For, by virtue of ergodicity, the empirical averages obtained along any simulation run must converge almost surely to the theoretical averages induced by the unique invariant distribution.

The former considerations suggest that numerical simulations are a sound way of obtaining suitable estimates of the long-run behavior of the system. Here, therefore, we pursue this route but complement it with a summary of the analytical discussion of the model undertaken by Marsili, Slanina & Vega-Redondo (2002). In this companion paper, we rely on the mean-field techniques widely used in statistical physics to formulate a suitable deterministic approximation of the dynamic behavior of the model that embodies many of its essential features. In combination, both approaches, numerical and analytical, jointly provide a rich perspective on the nature and intuition of the results.

In our discussion of the model, most of our efforts will be devoted to understanding the long-run relationships among the following variables:

- *network density*, as given by the average degree (or connectivity) of the nodes (i.e. players).
- *network cohesiveness*, as reflected by the average distance between the neighbors of every given node – that is, a measure of how closely connected are the neighbors of a typical player.
- *network span*, as embodied by the relative size of the largest components or, relatedly, by the number of distinct agents who fall within a some given radius of some typical player.
- *payoff performance*, as given by the average payoff earned per interaction (i.e. across all links).

As a “numerical confirmation” of ergodicity, the simulation paths indeed show convergence of the empirical average magnitudes of the above listed variables to a given value (i.e. their ergodic mean), even when the process starts from very different initial conditions. Our main concern then is to understand how these long-run summaries of process depend on the underlying data of environment (i.e. on the parameters of the model). Very succinctly, the main regularities can be summarized as follows.

- (a) The long-run density of the network depends negatively on ϵ , the extent of payoff volatility. So happens as well with payoff performance. In this sense, therefore, one obtains the intuitive conclusion that volatility is detrimental not only to the accumulation of social capital but also to the materialized payoffs.
- (b) As the environment’s volatility rises, the population “protects” from it by *endogenously* increasing the cohesiveness of the social network.

- (c) Whenever the society is able to sustain a dense social network, its architecture displays a high span, which turns out to be a crucial feature in enhancing the effectiveness of search (and thus adaptability).
- (d) The detrimental effect of volatility mentioned in (a) is, in fact, strongly mitigated by the strategic effects availed by the social network. If, by way of contrast, players refrained from using network effects to support cooperation (i.e. every bilateral IRPD game were strategically independent), the "harmful" effects of volatility would manifest themselves at much lower levels and in a substantially more abrupt fashion.

The above conclusions underscore the intuitive point that a "stable environment" may be an important requirement for successful accumulation of social capital. That is, only if agents' payoff conditions are not altered too fast can one hope that a dense social network of cooperation may be sustained over time. Indeed, we shall also see that the transition from a low- to a high-accumulation path may respond very drastically to just very slight changes of the underlying conditions. In a sense, this may be regarded as a stylized explanation of sudden "miracles" (or anti-miracles) in development or growth of the sort stressed, for example, by Lucas (1988).

Another important insight of the analysis concerns the way in which the architecture of social interaction responds *endogenously* to the underlying circumstances of the environment, thus attempting to offset its main detrimental implications. Somewhat schematically, the essential two features of the network architecture that bear on performance are *cohesiveness* and *span*. And, in accord with intuition, we observe that this trade-off moves in favor of cohesiveness (and away from span) as environmental volatility increases. It is worth advancing, however, that the process always tends to keep some balance between these two polar features, thus preserving – if volatility is not too intense – a "small-world"³ architecture for the (endogenous) social network.

The rest of the paper is organized as follows. Next, Section 2 presents the model – first, its static version in Subsection 2.1, then its dynamic counterpart in Subsection 2.2. The analysis of the model starts in Section 3 with the characterization of pairwise-stable networks. It proceeds with the dynamics in Section 4, which consists of two subsections: Subsection 4.1, that establishes some basic dynamic results (e.g the ergodicity of the process), and Subsection 4.2 that contains the bulk of our numerical analysis. In this latter subsection, the discussion starts with a benchmark scenario, followed by the consideration of a number of extensions and variants. Next, in Section 5 the mean-field analysis of the model is outlined, comparing its conclusions with the numerical simulations. Finally, in Section 6 the related literature is reviewed, while Section 7 offers some concluding remarks and a number of possible courses for future research.

³The notion of a *small world* – a network that displays both short distances and strong cohesiveness (or clusterization) – has been the object of booming research since the seminal work of Watts and Strogatz (1998).

2 Model

2.1 Statics

Let $N = \{1, 2, \dots, n\}$ be a finite population of agents who may interact through a collection of infinitely repeated games. For each pair of players who actually interact, $i, j \in N$, the stage game they play is an *idiosyncratic* Prisoner's Dilemma (PD) with payoff table

$i \backslash j$	C	D	(1)
C	ζ_{ij}	$\zeta_{ij} - \nu$	
D	$\zeta_{ij} + \nu$	0	

with $\nu > \zeta_{ij}(= \zeta_{ji}) > 0$. As customary, C and D will be labelled as ‘‘Cooperate’’ and ‘‘Defect,’’ respectively. The payoff obtained by both players if they jointly cooperate is denoted by ζ_{ij} . This payoff is ij -specific and, in the dynamic framework to be considered later on, it will change over time. For simplicity, the payoff of joint defection is normalized to zero, whereas in case of a unilateral defection the gain ν obtained by the defector over ζ_{ij} is made equal to the loss incurred her partner. To fix ideas, one possible interpretation of the situation is that the interaction between players i and j concerns the implementation of a joint project of total value $2\zeta_{ij}$, which is either divided equally if both behave ‘‘honestly’’ or allows the individual who unilaterally behaves dishonestly to appropriate an excess payoff of ν at the expense of the other party.

In general, of course, not all players interact among themselves. The specification of who does is given by the *social network*, as captured by a certain graph $g \subset N \times N$, where the nodes are identified with the players and $(i, i) \notin g$ for any $i \in N$. Naturally, this graph is undirected so that, for all $i, j \in N$, $(i, j) \in g \iff (j, i) \in g$. Typically, the more compact notation ij (or ji) will be used to denote the link between player i and j . Furthermore, we shall write $g - ij$ or $g + ij$ to represent the networks obtained from g by, respectively, adding or removing a link ij .

Given the prevailing network g , all pairs of players i and j such that $ij \in g$ are involved in an Infinitely Repeated Prisoner's Dilemma (IRPD) with idiosyncratic stage payoffs given by their respective ζ_{ij} – cf. (1). Each of these different IRPD are *choice independent*, in the sense that players' decisions in any one of them do not restrict the feasible behavior in others. They need *not* be, however, *strategically independent* since the behavior in some of them may be made contingent on the information of what has occurred in others.

A key feature of the present approach is that information on how players have behaved in the past diffuses through the social network only gradually. Specifically, it is assumed that, in every round of play of the IRPD played between two ‘‘neighbors’’ (i.e. connected players), each of them informs the other one of any worthwhile detail they might be informed at that point – most

crucially, of any deviation by other players from notionally prescribed behavior. In this manner, relevant strategic information “travels” along the network one step (or link) at a time. The essential implication of this assumption is that, in general, the architecture of the network has an important bearing on the extent of cooperation that the population as a whole can support in an incentive-compatible manner.

To simplify matters, let us suppose that players rely on *trigger strategies* (see Remark 2 for further discussion). In the present context, these strategies will be assumed to have the following format:

- (a) first, players choose, separately with each neighbor, whether to start their bilateral interaction by cooperation or defection;
- (b) subsequently, they immediately react to the news that one of her own neighbors has formerly behaved as a unilateral defector (with any player) by “punishing” her, i.e. by switching to irreversible defection in the corresponding bilateral IRPD played with that neighbor.

Now, under the assumption that players rely on trigger strategies, let us introduce the following convenient notation. Given some prevailing network g and any given player i , denote by

$$N_i \equiv \{j \in N : ij \in g\}$$

the set of her neighbors. For any one of these neighbors $j \in N_i$, let $g \overset{ij}{\sim} a$ stand for the situation (i.e. strategy profile) where cooperation is the initial action of *both* players for *every* link in g , except for the initial action of player i in link ij which is $a \in \{C, D\}$. This strategy profile univocally induces a corresponding infinite flow of *stage* payoffs for player i in her interaction with neighbor k :

$$\{\psi_{ik}^\tau(g \overset{ij}{\sim} a)\}_{\tau=0}^\infty \quad (2)$$

where $\psi_{ik}^\tau(g \overset{ij}{\sim} a)$ represents the stage payoff earned by player i with k at stage τ . Then, denoting by $\delta \in (0, 1)$ the common discount factor, the payoff flow (2) induces a (normalized) discounted payoff equal to

$$(1 - \delta) \sum_{\tau=0}^\infty \delta^\tau \psi_{ik}^\tau(g \overset{ij}{\sim} a). \quad (3)$$

Finally, the latter discounted payoffs can be aggregated to obtain the total payoff earned by player i over all her neighbors:

$$\pi_i(g \overset{ij}{\sim} a) \equiv \sum_{k \in N_i} \left\{ (1 - \delta) \sum_{\tau=0}^\infty \delta^\tau \psi_{ik}^\tau(g \overset{ij}{\sim} a) \right\}. \quad (4)$$

We are now in a position to define the notion of *Pairwise-Stable Network* (PSN). Informally, it simply consists of a network where, for every separate link,

the two players involved have incentives to use it for cooperation. Thus, this notion implicitly embodies the idea that, unless both of the players connected by each link can separately confirm its incentives for cooperation, that link will vanish. A formal definition of this idea follows.

Definition 1 *An undirected graph $g \subset N \times N$ is said to be a Pairwise-Stable Network if for all $ij \in g$,*

$$\pi_i(g \overset{ij}{\sim} C) \geq \pi_i(g \overset{ij}{\sim} D).$$

To end this subsection, the following remark clarifies certain interesting issues concerning the use of trigger strategies in the present context.

Remark 2 – Trigger strategies, maximal punishment, and perfection:

As formulated, the PSN concept directly embodies the assumption that players restrict to trigger strategies. This restriction was justified above on the grounds of simplicity. But in line with well-known results on the theory of repeated games (see Abreu(1988)), a further justification may be grounded in the fact that those strategies induce *maximal* punishments in the IRPD. They can be used, therefore, to support *any* incentive-compatible behavior – in other words, any outcome which could be supported by other, possibly more complex, strategies can be supported as well through those of a trigger type. In this sense, these strategies can be postulated without loss of generality if one is interested alone in characterizing incentive-compatible *behavior*.

Trigger strategies, however, raise in the present context a problem concerning issues of credibility (or perfection). In particular, it is not generally optimal for a player to punish a neighbor when news about the latter's deviant behavior arrives. For, by eschewing such a punishment, the player anticipates being able to sustain cooperation with her neighbor, an alternative which may well be preferred if such a reciprocal cooperation can indeed be induced. A natural way to address this conceptual problem is to modify the stability concept (and, correspondingly, enrich the set of admissible strategies) so that any potential defector is given the possibility of anticipating, and reacting optimally to, ensuing punishment. This, in turn, makes punishment of a (sophisticated) deviator optimal when the news to that effect arrive, even if the deviator has always cooperated. Such considerations give rise to what might be called a *Sophisticated Pairwise-Stable Network* (SPSN), a concept which displays the following advantages. Firstly, it has the virtue of guaranteeing stability with respect to the set of *all* possible strategies. Secondly, it is “perfect” in the sense of being supported by strategies that are optimal after every history – in particular, therefore, it renders punishment an optimal response in every case. In contrast, however, the main drawback of the SPSN concept is that, while it is substantially more complex to formalize and characterize than its non-sophisticated counterpart, its qualitative implications are essentially the same – see Remark 5. All combined, therefore, it seems well advised to focus on the PSN concept, for the sake of analytical simplicity and expositional clarity. \square

2.2 Dynamics

The equilibrium approach introduced so far represents an intermediate stage towards a fully dynamic description of the situation. To carry out the latter, however, we first need to propose in detail the law of motion that is to govern the social dynamics. This is the object of the present subsection.

Let time be modelled discretely, with $t = 0, 1, 2, \dots$ indexing the consecutive periods. At every t , each player $i \in N$ supports a certain number of links (what are called her direct links), while he may be also connected to other players through the latter's direct links – what may be called that player's passive links. Each of these links (active or passive) has a certain payoff potential associated to it, which is identified with the cooperation payoff in the stage PD game played by the two agents thus connected. Under the assumption that the support of a link is a resource-consuming activity, an important feature of the model is that players are limited to supporting a maximum number of links, say m , an exogenous parameter.

The above considerations suggest identifying the state of the system at (the beginning of) any given t by a list $\omega(t) \equiv [L_i(t)]_{i \in N}$ where each $L_i(t) = \{(j_1, \zeta_{ij_1}), (j_2, \zeta_{ij_2}), \dots, (j_r, \zeta_{ij_r})\}$ specifies the collection of individuals $j_1, j_2, \dots, j_r \in N \setminus \{i\}$ to whom player i supports a (direct) link, as well as the corresponding payoffs $\zeta_{ij_1}, \zeta_{ij_2}, \zeta_{ij_r} \in \mathbb{R}_+$ associated to each such link. As indicated, the cardinality r of the set $L_i(t)$ must not exceed m . Given any such state $\omega(t)$, the induced social network $g(t)$ is simply formed by declaring two individuals i and j linked if there is a (direct) link between them, in either direction. Thus, formally, we have:⁴

$$ij \in g(t) \Leftrightarrow (j, \cdot) \in L_i(t) \vee (i, \cdot) \in L_j(t),$$

the induced set of neighbors for any player i being denoted by $N_i(g(t))$.

The social dynamics defining the law of motion across consecutive states embodies three distinct sequential components: payoff update, link formation and search, removal of unstable links. I take up each of these in turn.

1. Payoff update

First, we suppose that the payoff of each link may be subject to a random update of its associated payoff. More precisely, with some independent probability ε , every link $(j, \zeta_{ij}) \in L_i(t)$ of each player i is transformed into the link (j, ζ'_{ij}) where ζ'_{ij} is drawn afresh from some non-negative real interval $[\underline{\zeta}, \bar{\zeta}]$ according to a stationary (and common) probability distribution with continuous density f_ζ . For future reference, denote by $\omega'(t) = [L'_i(t)]_{i \in N}$ the new set of links thus generated, with $g'(t)$ (equal, in fact, to $g(t)$) standing for the corresponding network.

⁴For simplicity, it will be assumed that (redundant) direct links in both directions do not exist between any pair of players. In fact, by assuming that there are no such links at the beginning of the process, the dynamics of link formation guarantee that they will never arise later on.

2. Link formation and search

In every period, every player $i \in N$ may enjoy two possible routes of search and consequent formation of fresh links: component-bound (“local”) search and unrestricted (“global”) search. Whereas the first route is conceived as the “ordinary” way of accessing new information (i.e. mediated by the social network), the second one is regarded as more extraordinary (and thus only occasional). Formally, we shall posit that, with independent probability p , the first option arises, whereas with probability $(1-p)q$ the second one occurs.⁵ I describe each of these alternative options in turn.

(2.a) Component-bound search

Given $g'(t)$, let $C_i(g'(t))$ represent the set of players who belong to the same component as i in $g'(t)$ – i.e. those players for whom there is a path in $g'(t)$ joining them to node i .⁶ When some player i receives a component-bound revision opportunity, let us suppose that she receives information on new ‘projects’ to be undertaken with each of the players in her component with whom she is not already interacting. More precisely, she is taken to observe a set of possible new links with the individuals in $C_i(g'(t)) \setminus N_i(g'(t))$, with its associated payoffs having been drawn afresh according to the probability density f_ζ . Let $S_i^0(t)$ stand for the set of these links. Out of this set, player i should be interested only in those who she perceives as pairwise-stable, given the prevailing state $\omega'(t)$ and the induced network $g'(t)$. Denote by $S_i^1(t) (\subset S_i^0(t))$ the set of those links which display such pairwise-stability,⁷ and let $l_i^* = (j^*, \zeta_{ij^*})$ be the particular link (or any one of them, randomly chosen, if there are several) which provides the highest payoff. Then, two cases must be distinguished.

- (i) If the cardinality $|L_i'(t)| < m$, and therefore player i is not currently saturating her link-supporting capability, then the new set of links supported by her is simply made equal to $L_i''(t) = L_i'(t) \cup \{(j^*, \zeta_{ij^*})\}$.
- (ii) If, instead, $|L_i'(t)| = m$, the new set $L_i''(t)$ is formed by the m links in $L_i'(t) \cup \{(j^*, \zeta_{ij^*})\}$ with highest payoffs.

⁵Therefore, with probability $(1-p)(1-q)$, any given player receives no revision opportunity whatsoever.

⁶Subsection 4.2.4 considers an arguably more realistic variant of the model where fresh links can only be formed with players who are close to them – i.e. in their own component, but within a short “search radius”. It turns out that the main gist of the analysis is unchanged by this modification.

⁷Formally, we adapt the notation used in Definition 1 and say that a new link of player i with some player j is perceived as pairwise stable, given $\omega'(t)$, when both

$$\begin{aligned} \pi_i'((g'(t) + ij) \overset{jj}{\rightsquigarrow} C) &\geq \pi_i'((g'(t) + ij) \overset{jj}{\rightsquigarrow} D) \\ \pi_j'((g'(t) + ij) \overset{ii}{\rightsquigarrow} C) &\geq \pi_j'((g'(t) + ij) \overset{ii}{\rightsquigarrow} D), \end{aligned}$$

where $\pi_i'()$ and $\pi_j'()$ reflect the payoffs displayed in $\omega'(t)$.

(2.b) *Unrestricted search*

If some player i receives a revision opportunity through global search, she gets to observe the possibility of forming a new link $l_i^\diamond = (j^\diamond, \zeta_{ij^\diamond})$, where j is *any* player in $N \setminus N_i(g'(t))$ and, again, ζ_{ij^\diamond} is drawn according to the probability density f_ζ . Suppose that l_i^\diamond is pairwise-stable, given $\omega'(t)$. Then, as above, the set $L_i''(t)$ of new links of player i is simply made equal to $L_i'(t) \cup \{(j^\diamond, \zeta_{ij^\diamond})\}$ if $|L_i'(t)| < m$, whereas it is identified with the subset of m links with the highest payoff otherwise. Finally, if l_i^\diamond is not perceived as pairwise-stable by player i , it is simply ignored and no actual revision occurs.

3. Removal of pairwise-unstable links

Let $\omega''(t) = [L_i''(t)]_{i \in N}$, with g'' denoting the induced network and $\pi_i''(\cdot)$ the payoffs associated to the corresponding payoff profile $[\zeta_{ij}'']_{ij \in g''(t)}$. (For notational simplicity, I dispense with the time index.) Then, for every link $ij \in g''$, let the players i and j involved in the link evaluate whether both of the following incentive-compatibility conditions hold:

$$\begin{aligned} \pi_i''(g'' \overset{ij}{\sim} C) &\geq \pi_i''(g'' \overset{ij}{\sim} D) \\ \pi_j''(g'' \overset{ji}{\sim} C) &\geq \pi_j''(g'' \overset{ji}{\sim} D). \end{aligned}$$

If either of these conditions is violated, the link ij is judged unstable by the players and thus is removed.⁸ Once such a check of pairwise-stability has been completed for all links in g'' , let $\omega'''(t)$ refer to the resulting state where only the links that qualified as pairwise-stable remain. This state is then carried over to the next period, by making $\omega(t+1) = \omega'''(t)$.

3 Static analysis: characterization of Pairwise-Stability

As we did in the presentation of the model, it is useful to start our discussion with an equilibrium (thus static) approach, subsequently enriching matters with a full-fledged dynamic analysis. Proposition 3 below initiates this course by providing an intuitive characterization of Pairwise-Stable Networks. As suggested in our informal discussion, this characterization crucially hinges upon a certain measure of cohesiveness of the network, as given by suitably defined geodesic distances concerning the neighbors of the different players. Formally, for any given player $i \in N$, and any two of her neighbors $j, k \in N_i$, we define the *i-excluding distance* between j and k , denoted by $d^i(j, k)$, as the length of the shortest path joining j and k which do not involve player i . The interpretation

⁸Note that, in the spirit of Definition 1, we are making the assumption that each player evaluates separately the stability of each of her different links. Moreover, for the sake of tractability, we also make the assumption that all considerations are made with respect to the same network g'' prevailing at the start of the link-removal operation.

of this distance is straightforward: it is the number of steps (and therefore periods, in the repeated game) which are required for an information held by j (or k) to reach k (or j) without the concurrence of player i . As usual, it is postulated that $d^i(j, j) = 0$ for any $j \in N_i$, while if no i -excluding path exists between k and j it will be convenient to posit that $d^i(j, k) = \infty$.

Proposition 3 Consider any network $g \subset N \times N$ and let $[\zeta_{ij}]_{ij \in g}$ stand for the possible cooperation payoffs that can be earned for each of its links. Then, g is a Pairwise-Stable Network (PSN) if, and only if, for all $ij \in g$:

$$\frac{1}{2} \sum_{k \in N_i} \left(\zeta_{ik} + \frac{1-\delta}{\delta} \nu \right) \delta^{d^i(j,k)} \geq \frac{1-\delta}{\delta} \nu \quad (5)$$

Proof. Consider any link $ij \in g$ and focus, for concreteness, on player i . Pairwise-stability of this link requires that player i has incentives to cooperate with j under the threat that, if she were to do otherwise, all his neighbors $k \in N_i$ will switch to defection once they learn about it – an event that, for each of them, occurs $d^i(j, k)$ periods after the contemplated defection takes place.

If player i cooperates with j , she anticipates an intertemporal payoff:⁹

$$\pi_i(g \stackrel{ij}{\sim} C) = \sum_{k \in N_i} \zeta_{ik}.$$

Instead, if player i defects unilaterally upon j , her anticipated payoff is:

$$\begin{aligned} \pi_i(g \stackrel{ij}{\sim} D) &= (1-\delta)(\zeta_{ij} + \nu) \\ &+ \sum_{k \in N_i \setminus \{j\}} \left\{ \left[\sum_{s=0}^{d^i(j,k)-1} (1-\delta)\delta^s \zeta_{ik} \right] + (1-\delta)\delta^{d^i(j,k)} (\zeta_{ik} - \nu) \right\} \\ &= (1-\delta)(\zeta_{ij} + \nu) + \sum_{k \in N_i \setminus \{j\}} \left[(1-\delta)^{d^i(j,k)+1} \zeta_{ik} - (1-\delta)\delta^{d^i(j,k)} \nu \right]. \end{aligned}$$

Therefore, the stability condition

$$\pi_i(g \stackrel{ij}{\sim} C) \geq \pi_i(g \stackrel{ij}{\sim} D)$$

can be written as follows:

$$\delta \zeta_{ij} + \sum_{k \in N_i \setminus \{j\}} \delta^{d^i(j,k)} \delta \zeta_{ik} \geq (1-\delta)\nu - \sum_{k \in N_i \setminus \{j\}} (1-\delta)\delta^{d^i(j,k)} \nu$$

which is readily seen to be equivalent to (5).

⁹Recall that stage payoffs are normalized by the factor $(1-\delta)$.

■

Remark 4 *Network stability in the absence of network effects:*

If players did not rely on network (population-wide) effects in their strategic considerations, cooperation could be supported through any particular link ij if, and only if, it could be done bilaterally. That is, if

$$\zeta_{ij} \geq \frac{1-\delta}{\delta}\nu. \quad (6)$$

Contrasting (5) with (6), it immediately follows that, naturally, network-based effects can only help in supporting cooperation, i.e. the latter condition always implies the former. In general, the strength of those effects positively depends on the number of neighbors a player has, how valuable these are (i.e. the opportunity cost of spoiling cooperation with them), and their relevant distance (which impinges on the delay affecting the aforementioned costs).

It should be emphasized, however, that even in the absence of network strategic effects, the architecture of players' interaction may have an important bearing on the social dynamics through its effect on search and innovation. This, in fact, will be confirmed by the dynamic analysis of such a scenario (cf. Subsection 4.2.2), where some of the regularities displayed by the basic model still arise. For example, payoff volatility continues to have a negative impact on long-run performance, even though many of the important features that go along with this phenomenon (e.g. network cohesiveness, payoff performance, etc.) are sharply affected by the absence of network-mediated feedback on players' behavior. \square

Remark 5 *Network stability with higher player sophistication:*

In Remark 2, we discussed the possibility of allowing players a superior degree of sophistication that would allow them to anticipate the stage at which others would punish her for a deviation, thus reacting optimally to it through non-trigger strategies. Space limitations prevent us from developing in detail the implications of the stability concept which would reflect these considerations and that we called *Sophisticated PSN*. The interested reader, however, may verify that, as a counterpart of (5), the condition characterizing pairwise-stability in that case would read as follows:

$$\zeta_{ij} - \frac{1-\delta}{\delta}\nu \geq \min\{0, - \sum_{k \in N_i \setminus \{j\}} (\zeta_{ik} - \frac{1-\delta}{\delta}\nu) \delta^{d^i(j,k)-1}\}. \quad (7)$$

For the sake of a more transparent comparison, the above condition can be contrasted with the following rewriting of (5):

$$\zeta_{ij} - \frac{1-\delta}{\delta}\nu \geq - \sum_{k \in N_i \setminus \{j\}} (\zeta_{ik} + \frac{1-\delta}{\delta}\nu) \delta^{d^i(j,k)-1}.$$

One then observes that similar qualitative considerations arise in both cases – i.e. the number, payoff value, and relevant distances of neighbors continue to

be the key factors involved. In (7), however, the higher sophistication assumed on the part of players (which allows them a preemptive reaction to punishment) leads to weaker deterrence against deviations. This, in turn, narrows – but certainly does not destroy – the range of circumstances in which a sufficiently cohesive social network enlarges significantly the cooperation possibilities. \boxtimes

4 Network dynamics

4.1 Ergodicity and other preliminary results

Suppose that no fresh links were ever formed through search (i.e. $p = q = 0$) and prevailing payoffs were not subject to update ($\varepsilon = 0$). Then, the resulting deterministic dynamics would reduce to a mere chain of link removals, as the links which are deemed unstable are being discontinued by the agents involved. This restricted dynamics would obviously lead to a stationary situation, possibly an empty network. In this sense, it may be regarded as providing a simple dynamic foundation of the PSN concept introduced in Definition 1. But our main interest, of course, is geared towards the study of the full-fledged stochastic dynamics where link removal is countered by search and link formation, as well as the underlying payoff conditions are subject to occasional change. Only in this case, with $p, q, \varepsilon > 0$, the full richness of our model is suitably displayed.

A first basic step to be undertaken in the analysis of such stochastic dynamics is the establishment of conditions under which the induced Markov process can be ensured ergodic. As advanced, such ergodicity will provide a theoretical basis for the later use of numerical simulations in elucidating the long-run behavior of the system. To state matters formally, recall that $f_\zeta(\cdot)$ stands for the probability density which governs every fresh draw of payoffs, whose support is given by a non-negative interval $[\underline{\zeta}, \bar{\zeta}]$. Further remember from (6) that $\frac{1-\delta}{\delta}\nu$ is the threshold to be exceeded by any cooperation payoff ζ_{ij} which is to be supportable by players i and j *without* resorting to network effects (i.e. bilaterally). In this terms, the following result states that ergodicity is guaranteed as long as, for any payoff value redrawn afresh, there is some prior uncertainty as to its “supportability” in the absence of network effects.

Proposition 6 *Assume $\underline{\zeta} < \frac{1-\delta}{\delta}\nu < \bar{\zeta}$ and $\varepsilon, q > 0$. Then, provided the population N is large enough, the social dynamics described in Subsection 2.2 is governed by an ergodic stochastic process.*

Proof. The induced process is clearly aperiodic. Therefore, to establish the desired conclusion, it is enough to show that there is some particular state to which there is positive probability of returning, from any other state, in some finite number of steps. In the argument, the state ω^e where there are no links established between players will play this recurrent role.

The first point to note is that, if n is large enough, there is positive probability that the network might be eventually divided into two or more disjoint components. To see how this could occur, suppose that n is even and let the

population be partitioned into two disjoint subsets, say $N^1 = \{1, 2, \dots, \frac{n}{2}\}$ and $N^2 = \{\frac{n}{2} + 1, \frac{n}{2} + 2, \dots, n\}$. Then, revision opportunities may arise (possibly through unrestricted search – recall (2.b) in Subsection 2.2) so that players in N^1 obtain payoff options ζ_{ij} with players $j \in N^2$ satisfying:

$$\bar{\zeta} > \zeta_{ij} > \max\left\{\frac{1-\delta}{\delta}\nu, \max_{k \in N^2} \zeta_{ik}\right\},$$

and reciprocally for players in N^2 . These revision opportunities induce pairwise-stable and payoff-improving links for each player in N^u over those that could be supported with agents in the complementary set N^v ($u, v = 1, 2, u \neq v$). Therefore, if n is large enough (in particular, it is enough that $n \geq 2(2m+1)$), all links across N^1 and N^2 would eventually be removed.

Thus suppose that players are divided into such disjoint components, and let each player in N^u in turn receive an unrestricted revision opportunity with some player in N^v ($v \neq u$) whose associated payoff is higher than any prevailing one (but lower than $\bar{\zeta}$). Then, the corresponding link must be formed (removing one of the pre-existing ones), since that link will be perceived as pairwise stable. Now suppose that, subsequently, the link is subject to a payoff update, with the consequence that its payoff is lowered below $\frac{1-\delta}{\delta}\nu$. This leads to the link being removed, since it is the only link which currently connects the sets N^1 and N^2 and, therefore, players cannot rely on network effects to support cooperation with it. By proceeding in this fashion with all players in turn as needed, it is clear that the process would reach the empty network. By construction, the chain of steps involved is finite and displays positive probability, thus completing the proof. ■

The former result clarifies that some payoff volatility (i.e. $\varepsilon > 0$) is generally needed to achieve ergodicity of the process. But the role played by ε in the analysis is not merely technical (i.e. to ensure ergodicity), but represents one of the key factors motivating both the the model and the questions asked. It is of central interest, therefore, to gain insight into the role it plays in the long-run dynamics of the process, understanding its interplay with other features of the model, e.g. payoffs. At this point of the discussion, a useful first step in this endeavor is afforded by the following result, which maintains all the hypotheses of Proposition 6 except for assuming that $\varepsilon = 0$.

Proposition 7 *Assume $\underline{\zeta} < \frac{1-\delta}{\delta}\nu < \bar{\zeta}$ and $q > 0$. Then, if $\varepsilon = 0$, the social dynamics leads almost surely to a path where the network reaches the maximum connectivity and the induced total payoff (aggregated over the whole population) converges to its maximum value $4mn\bar{\zeta}$.*

Proof. To establish the desired conclusion, the key role is played by the following two observations:

- (1) Consider any $\eta > 0$ such that $\bar{\zeta} - \eta > \frac{1-\delta}{\delta}\nu$. Then, since the density $f_{\zeta}(\cdot)$ is assumed continuous on its support $[\underline{\zeta}, \bar{\zeta}]$ and revision opportunities are

independent across players and time, the following conclusion applies. For all $\theta > 0$, there is some T such that if $t \geq T$, there is probability no lower than $1 - \theta$ that every player i has received (in preceding periods $\tau < t$) at least m link formation opportunities with associated payoffs $\zeta_{ij} > \bar{\zeta} - \eta$.

- (2) Any of the link opportunities described in (1) are pairwise stable. Therefore, choosing θ and T as above, there is probability no lower than $1 - \theta$ that, if $t \geq T$, every player i is supporting m links at t (the maximum number), all of them with associated payoffs no lower than $\bar{\zeta} - \eta$.

Thus, since η and θ in (1)-(2) can be chosen arbitrarily small, the desired conclusion immediately follows. ■

The previous result indicates that, in the absence of payoff volatility ($\varepsilon = 0$), the accumulation of social capital eventually reaches its maximum level of fruition. This in turn highlights the importance of such volatility in having other features of the dynamics – e.g. the evolving network architecture – enjoy some influence in shaping the long-run performance.

To further clarify the implications of Proposition 7, note that, if $\varepsilon = 0$, the process is no longer guaranteed to be ergodic and, therefore, one can no longer speak of a uniquely induced behavior in the long-run. Thus, in principle, the conclusion stated in the above proposition may arise in conjunction with long-run dependence of initial conditions concerning some other underlying features of the situation – e.g. the specifics of the network architecture. However, the fact that the invariant distribution of any Markov process is always upper hemi-continuous in its parameters implies that, for low but positive ε , the (then unique) long run behavior of the process must continue to display, approximately, the same features established by the result.

Finally, it is worth emphasizing that a crucial assumption in Proposition 7 is that $\frac{1-\delta}{\delta}\nu < \bar{\zeta}$. This amounts to saying that, with some positive probability, a fresh payoff draw may give rise to a potential a link which could be supported without resorting to network effects. To underscore the importance of this assumption, it may be useful to introduce the following observation.

Remark 8 *Network dynamics under stringent payoff conditions:*

Suppose that, unlike what has been assumed above, $\bar{\zeta} < \frac{1-\delta}{\delta}\nu$. Then, it is straightforward to see, that for any $\varepsilon \geq 0$ (thus, in particular, even in no payoff volatility exists), if the process starts with an empty network – i.e. what was labelled state ω^e in the proof of Proposition 6 – the network remains empty forever after. The simple reason for this is that, under these circumstances, no potential link is ever perceived as pairwise stable. Therefore, no link is ever formed, making the state ω^e an absorbing state. In contrast with Proposition 7, this points to the fact that the the long-run behavior of the system displays some sharp qualitative dependence on the payoff conditions. ☒

Combining Proposition 7 and Remark 8, a rich interplay is suggested between the exogenous conditions that determine payoffs and the extent of their

volatility. To explore this relationship, possible trade-offs, and their implications for the long-run performance of the process is the object of the next sections.

4.2 Numerical analysis

Motivated by the considerations explained in the previous Subsection, we now undertake an analysis of the model based on numerical simulations. First, the analysis will focus on a benchmark scenario where payoff conditions are set at the interesting “low edge” of the relevant region and players are allowed to take full advantage of network effects in supporting cooperation. Then, in subsequent subsections, we shall turn to exploring how the analysis is affected if players do not rely on network effects or the payoff conditions change – e.g. become more stringent.

4.2.1 A benchmark scenario

Consider an environment defined by the following parameter values. There are $100(= n)$ individuals involved in repeated interaction with a common discount rate $\delta = 3/4$ and a stage PD game with payoffs as given by (1) for $\nu = 4$. The cooperation payoffs ζ_{ij} are drawn, randomly and independently, according to a uniform distribution over the interval $[\underline{\zeta}, \bar{\zeta}] = [0.4, 1.4]$. Thus, as required by Proposition 6, the threshold $\frac{\nu(1-\delta)}{\delta} = \frac{4}{3}$ which marks the possibility of supporting cooperation bilaterally (cf. (6)) belongs to the payoff support. More specifically, it may be computed that any new payoff draw exceeds that threshold with an ex-ante probability of $1/15$.

The rate at which individuals receive search and revision opportunities is chosen to be $p = 0.1$ if they are component-bound, and $q = 0.01$ if they are unrestricted. On the other hand, the maximum number of links that any given player can actively support is set to $m = 2$ (therefore, the maximum average degree one may ever observe is equal to four).

The typical simulation runs will stretch for half million periods ($T = 5 \times 10^5$), a time horizon which always proves sufficient to attain long-run convergence in a clear-cut fashion. Our analysis of these simulations will focus on the key variables used in our earlier discussion: *network density* (average node degree), *network cohesiveness* (average neighbor distance), *network span* (relative size of the largest components), and *payoff performance* (average payoff earned per link). The main concern is to determine their respective long-run magnitudes, understanding how these long-run values depend on the essential details of the environment. As advanced, the key parameter in the analysis turns out to be the volatility rate ε , which will thus occupy a central role in our discussion. In a later subsection, however, we shall also discuss briefly how the behavior of the model is affected by changes in other parameters of the model such as, for example, the width of the payoff support.

For the scenario just described, we start our discussion with Figures 1 and 2, which show sample paths for the average degree of the network under three different volatility rates (i.e. different payoff-update probabilities),

$\varepsilon = 0.01, 0.07, 0.013$. Whereas the first diagram covers the complete time horizon of the simulation, i.e. half million periods, the second one provides a “zoom” restricted to the first 30000 periods. In both cases, the paths start with an empty network.

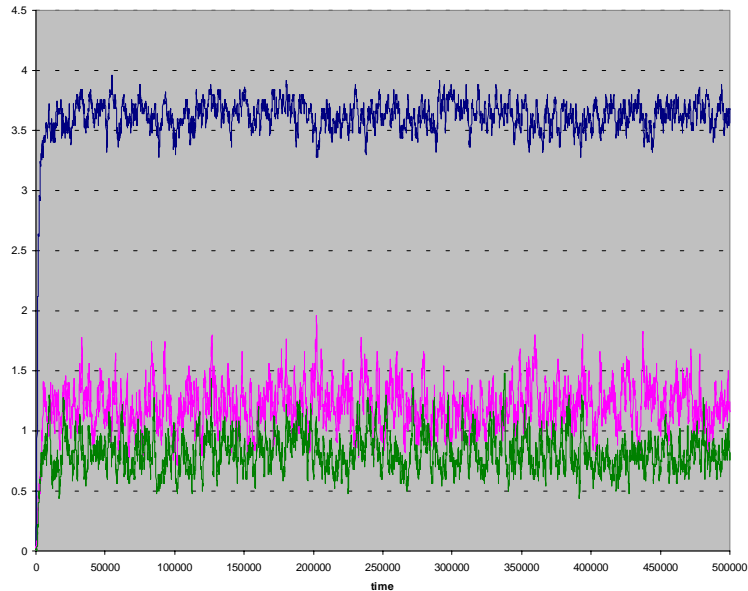


Figure 1: Average degree, full time horizon.

These diagrams suggest a rather fast convergence of the empirical averages, a feature that is specifically confirmed by Figure 3. Indeed, given the established ergodicity of the process, we know that the empirical averages observed along any simulation run must converge, a.s., to the corresponding magnitudes induced by the unique invariant distribution of the process. The fact that this convergence is quite fast and independent of initial conditions is further confirmed in Figure 4. In this latter figure, it is shown how, for each of the three volatility rates considered, the empirical averages converge quite rapidly to a common value, even if the alternative paths starting from *polar initial conditions*. More specifically, one of these initial conditions correspond to the empty network, as considered before. On the other hand, the polar case is given by paths that start at configurations where all players are connected to their neighbors through their maximum number of two active links.¹⁰

Analogous diagrams are displayed in Figures 5, 6, and 7 for the other three variables of interest: neighbor distance, largest-component share, and average

¹⁰To be concrete, what is done in this case in order to construct the initial conditions is to have every player i support random-payoff links to players $i+1$ and $i+2$, where these indices are conceived “modulo n ”.

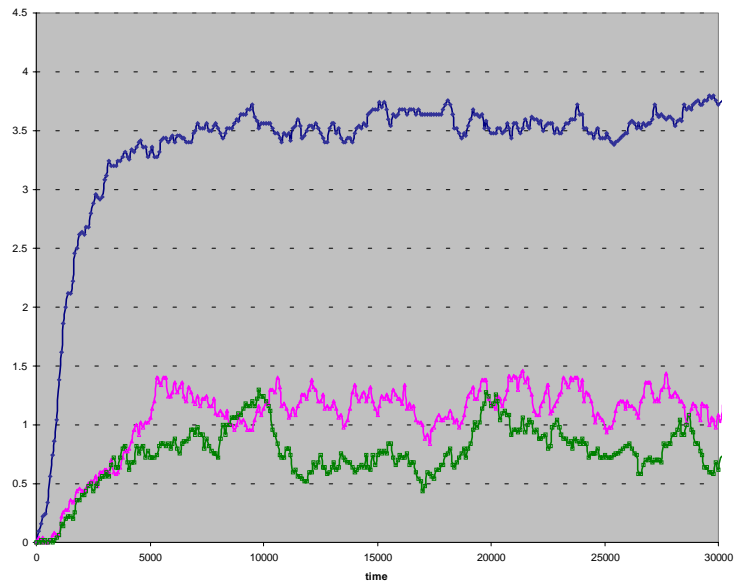


Figure 2: Average degree, short time horizon

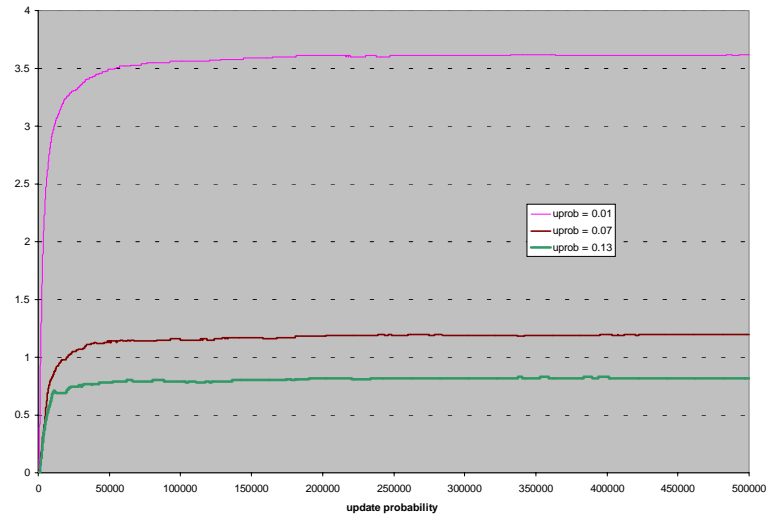


Figure 3: Empirical average degree

payoff. In every case we continue to observe that, for all three volatility rates considered, the empirical magnitudes converge rapidly to a common value even when starting from polar initial conditions. This shows again that the numerical simulations are an effective way for approximating the long-run values induced by the unique invariant distribution of the process.

The procedure outlined has been used to find approximate values for the long-run magnitudes of the variables of interest under a wide range of volatility rates ε . Our main interest in this respect is to understand the effect of ε on those long-run values. This is depicted in Figures 8-13 for each of the four cases: degree, neighbor distance, largest-component share, and link payoffs. Next, I discuss in turn the main findings obtained for each of them.

First, Figure 8 shows that payoff volatility has a negative effect on network connectivity. As the update probability ε rises, the density of the social network, as measured by the average node degree, is seen to fall quite significantly. Intuitively, this is a consequence of the fact that, as volatility grows, the pairwise stability of former links tends to deteriorate. This phenomenon may be regarded as a sort of depreciation in formerly accumulated social capital; or, in other words, as a reflection of the negative drift imposed on the value of existing links by the process of payoff update. For future reference (cf. Subsection 4.2.2), it is worth noting at this point that the decreasing function describing this effect is convex, the fall being quite steep for low values of ε but flattening significantly at higher values.

Figure 9 depicts what is possibly one of the most remarkable regularities found in the analysis. It pertains to the fact that, as ε grows, the social network endogenously adapts to this turn of events by becoming more *cohesive*. This is in line with what has been learned (say, in Proposition 3 concerning the PSN notion) about the role played by network effects in enhancing the incentives for cooperation. It is interesting to find that the network dynamics is responsive to these considerations, adapting endogenously to take advantage of them. As a graphic illustration of this phenomenon, Figures 10 and 11 display the networks prevailing at the end of the time horizon for two very different scenarios: one where payoff volatility is low ($\varepsilon = 0.02$) and another where it is relatively high ($\varepsilon = 0.12$). In line with our previous discussion, it is observed that in the first case there are a significant number of nodes that are loose ends of network paths – a sign of relatively low cohesiveness – while no such state of affairs is found in the second case (where, moreover, neighbors are connected on average by shorter paths).¹¹ Finally, it should be emphasized that the effects at work in this respect are strong enough to offset the important opposite forces which decrease general connectivity as ε grows (cf. Figure 13). For, *ceteris paribus*, such a decrease in node connectivity induces a marked tendency for *any* measure of graph-theoretic distance (even neighbor distance) to *rise*. A clearer understanding of this point will be gained in Subsection 4.2.2, where we consider a context with no network effects (and therefore no strategic benefit for cohesiveness).

¹¹More specifically, it may be computed that the neighbor distance of the first network (for

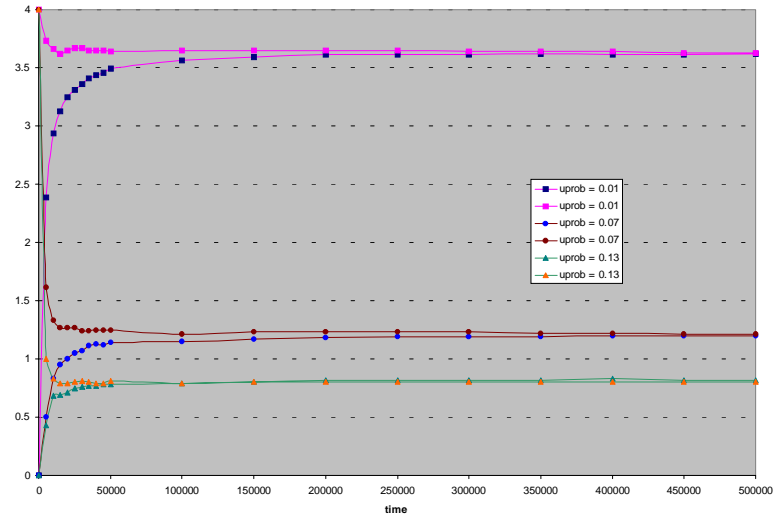


Figure 4: Empirical average degree, polar initial conditions

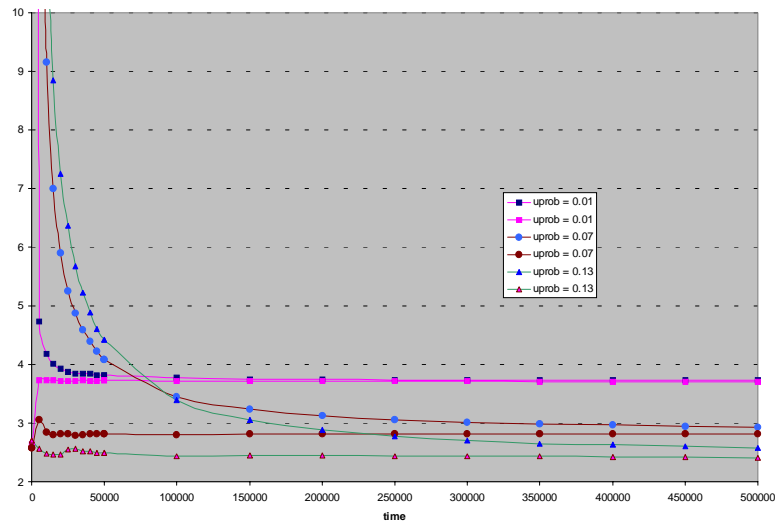


Figure 5: Empirical average neighbor distance, polar initial conditions

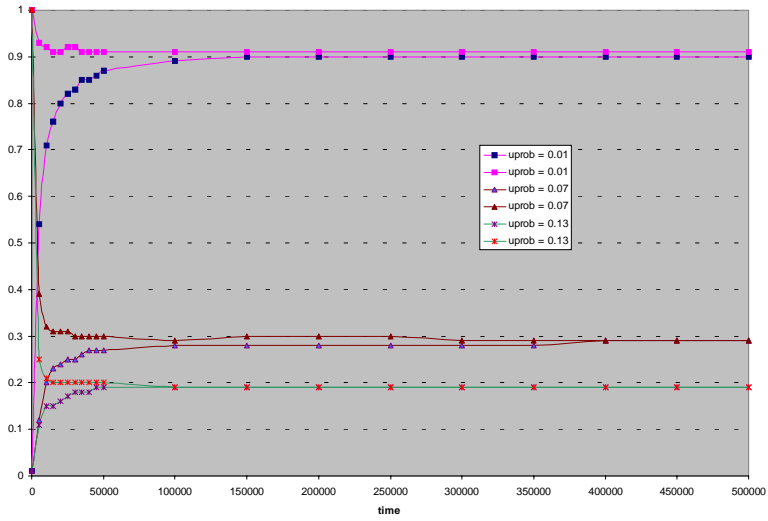


Figure 6: Empirical share of the largest component, polar initial conditions

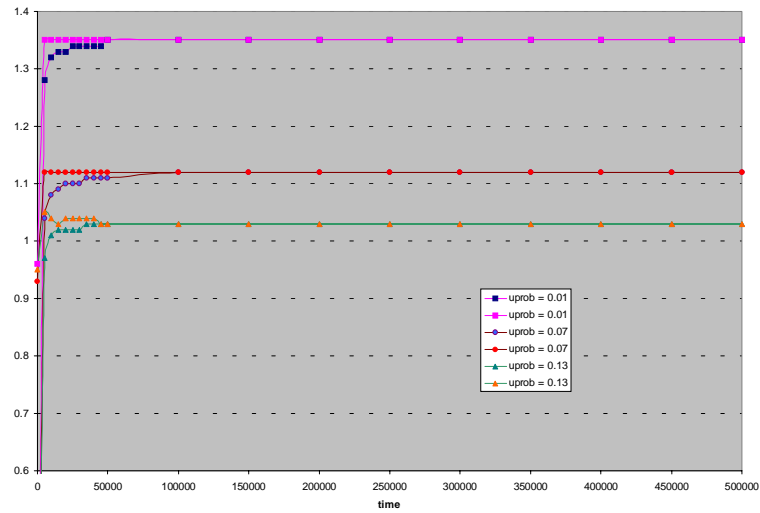


Figure 7: Empirical average payoff, polar initial conditions

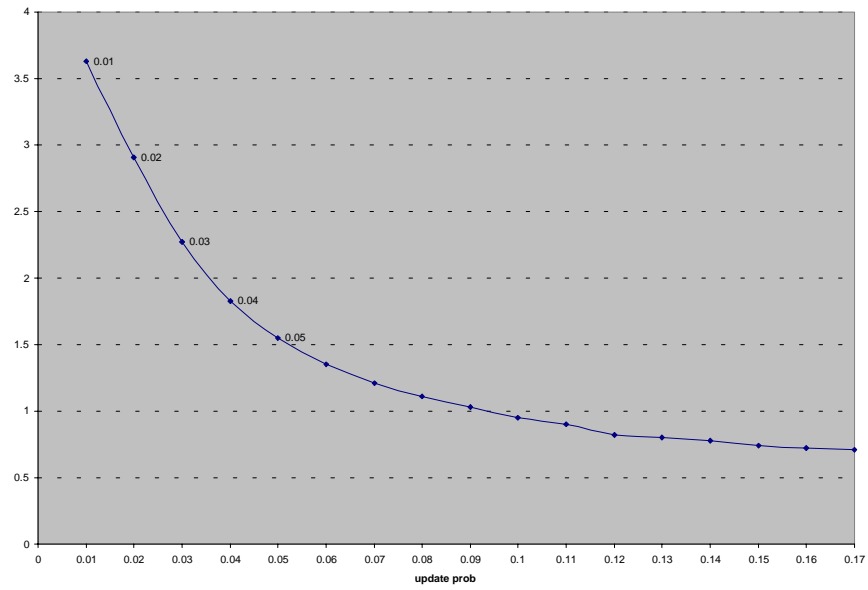


Figure 8: Payoff volatility and long-run average degree

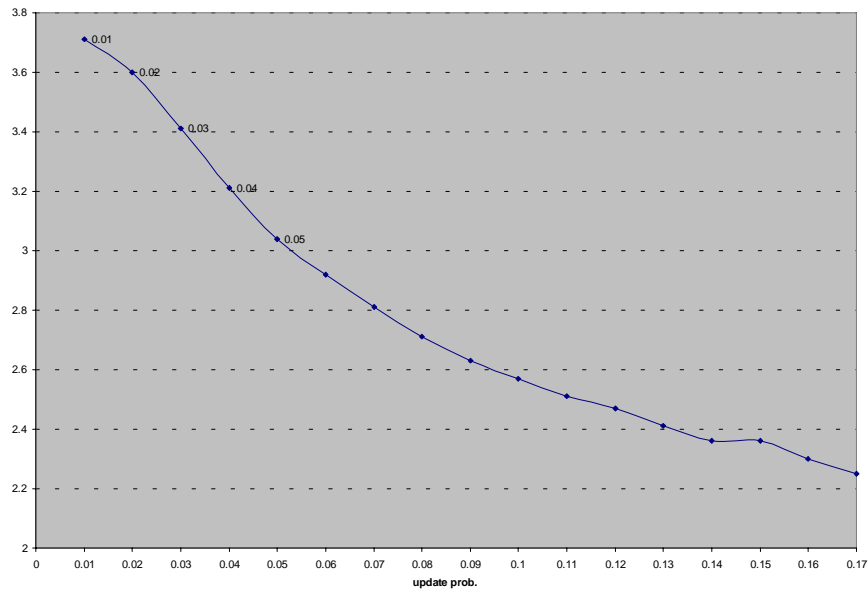


Figure 9: Payoff volatility and long-run average neighbor distance

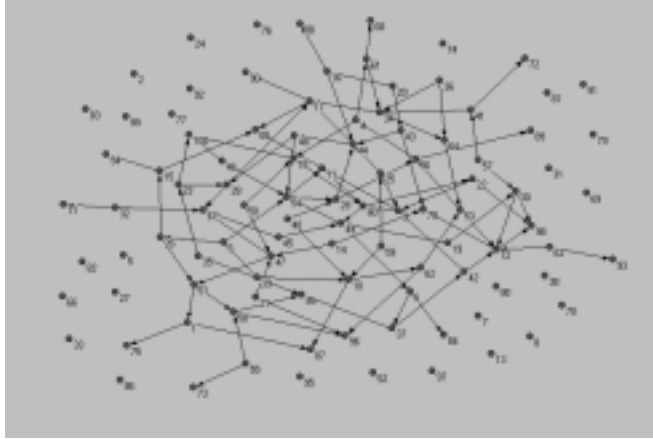


Figure 10: Social network after $5 \cdot 10^5$ periods, $\varepsilon = 0.02$. Any directed arrow joining two nodes represents a link, with its origin being the agent who supports that link. Nodes which have no incoming or outgoing link are fully disconnected.

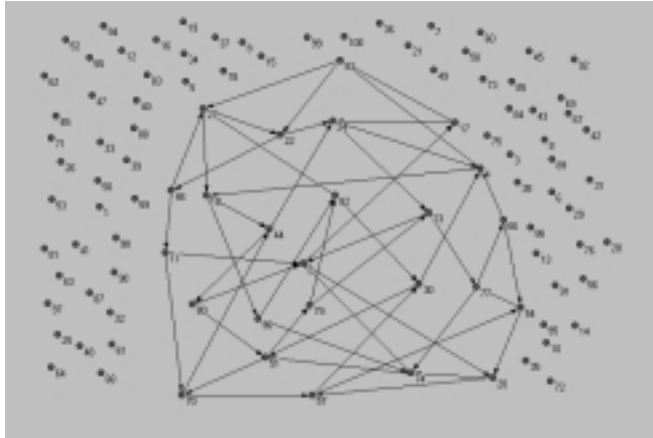


Figure 11: Social network after $5 \cdot 10^5$ periods, $\varepsilon = 0.12$. The same graphical conventions as above apply.

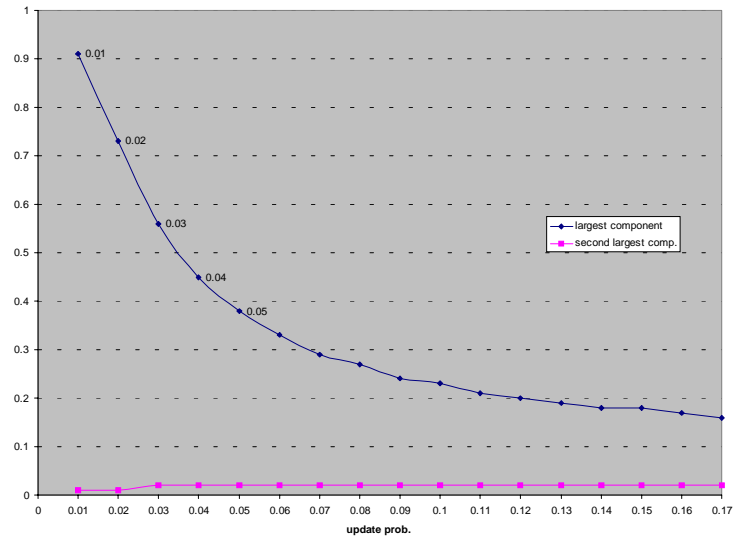


Figure 12: Payoff volatility and long-run share of two largest components

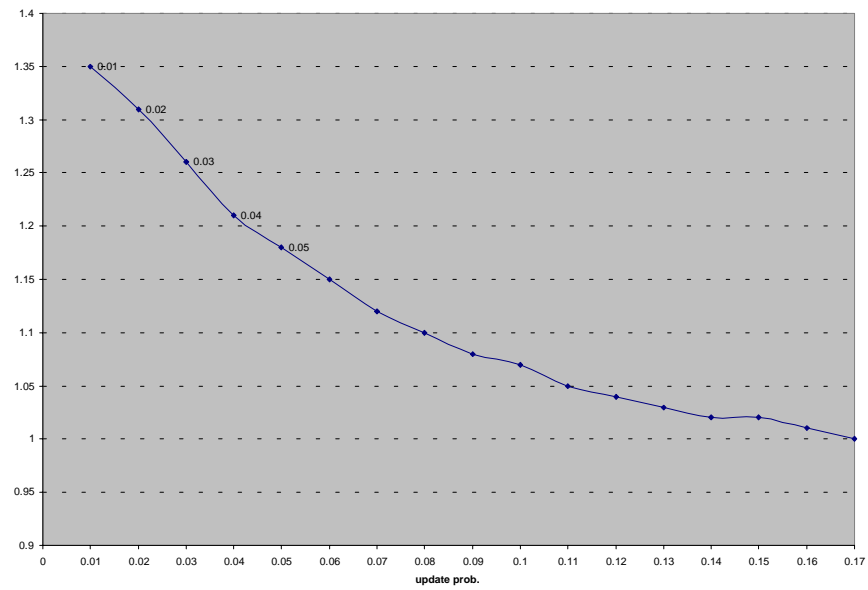


Figure 13: Payoff volatility and long-run average payoff

Figure 12 depicts the effect of payoff volatility on the size of the largest two components. The first interesting observation is that, independently of ε , most of the connected players belong to a *single* major component, the second-largest component remaining very small throughout. This conclusion is somewhat reminiscent of the well known results of Theory of Random Graphs which assert that, beyond a certain “connectivity threshold”, there arises a *single* large component in the graph. (These mathematical results, however, do not seem readily applicable here since, in our case, the formation of the social network is very much the result of a non-random mechanism.) On the other hand, concerning the effect of ε on the relative size of the largest component, an expected negative dependence is observed, largely a reflection of the analogous behavior displayed by the long-run average degree.

Finally, we turn to Figure 13, which centers on the effect of payoff volatility on the average payoff earned per interaction (or link). The evidence gathered now serves as an interesting complement to that displayed in Figure 8. It shows that, as ε grows, the fall in the average payoff earned *per link* reinforces the negative consequence induced by the formerly discussed decrease in the number of links (or games played). This latter effect was explained as an embodiment of social-capital depreciation. In contrast, the present one (a decrease in the payoff per link) should be largely understood as a consequence of the negative effects on search imposed by low connectivity and a small network component. The latter has detrimental implications for the bulk of players’ search activities (which are component-bound), in turn deteriorating the outcome of this search (the payoffs attained), even when such search is successful (i.e. happens to find a supportable link).

To sum up, the numerical analysis contributes the following insights on the factors that underlie a robust process of social-capital accumulation. First and foremost, we find that some environmental stability is required, and indeed the underlying network dynamics always tries to respond to this feature by adapting endogenously to the extent of prevailing volatility – in particular, by modulating its cohesiveness accordingly. When the process is successful, the increasing span of the social network is also exploited in order to enhance the scope of search. This allows players to maintain high levels of overall payoff performance and provides them with some important adaptability. Indeed, it is precisely this adaptability that, in a changing environment, is crucial to sustaining a stable density of interaction.

Overall, therefore, we find that *positive feedback effects* play a key role in the process (i.e. the network buildup reinforces itself) since a denser network both impairs strategic opportunism and enhances search. A useful course to take for improving our understanding of such positive feedback is to study how the dynamics would fare if it were either partially absent or somehow weaker. This is the objective of the following three subsections. The first one focuses on a context where no network effects are assumed to play any role in strategic considerations and therefore players’ cooperation has to be supported just

$\varepsilon = 0.02$) is 3.6 whereas that of the second (for $\varepsilon = 0.12$) is 2.7.

bilaterally. Then, the subsequent subsection explores whether a deterioration of players' payoffs (specifically, a downward shift in the payoff support) has a profound and lasting effect on the build up of social capital. Finally, the third subsection studies the long-run effects of restricting component-bound search to lie within a relatively short radius of the agent enjoying a revision opportunity.

4.2.2 Network-free strategic behavior

Suppose that, as described in Remark 4, players' strategic behavior no longer is influenced by network (i.e. inter-neighbor) effects. This may be understood as reflecting a different norm or convention used in the society – one where players react to each of their partners only according to the information gathered on their corresponding bilateral play. Then, strategically speaking, every pairwise interaction is to be regarded as strategically independent of any other. The role of the network consequently reduces to defining the channels through which information diffuses in the population, a nevertheless important role that should still yield interesting dynamic implications. As advanced, contrasting these implications with those obtained in our benchmark scenario, we hope to gain further insights on the workings of the model.

Figures 14-17 summarize the long-run effect of payoff volatility on each of the four variables that guide our analysis: degree, neighbor distance, largest-component share, and link payoffs. (The parameter values are all identical to those considered in the benchmark scenario.)¹² Comparing the long-run behavior displayed with its counterpart for the basic model (cf. Figures 8-13), the main points of contrast can be summarized as follows:

1. The bite on social performance caused by increased volatility manifests itself at much lower rates than in the original scenario. Specifically, both the average degree and the network span (cf. Figures 14 and 16) start to face significant downward adjustments at starting values of ε that are *one order of magnitude* smaller than before. A similar pattern is displayed in Figure 17 for the average payoff per link. In this case, however, the downward movement is more limited than in the benchmark scenario since, naturally, it is bounded below by the bilateral-cooperation threshold $\frac{\nu(1-\delta)}{\delta} = \frac{4}{3}$. Of course, such a more limited range of variation on the materialized payoffs is achieved at the expense of a much more pronounced reduction on the number of links as ε rises.
2. An additional, and very interesting, difference with the original framework concerns the abruptness now observed in the transition from a high- to a low-performance regime. In the absence of network strategic effects, this transition takes place rather sharply, at least in two complementary senses. On the one hand, the full change is essentially completed within a range

¹²By relying on ideas analogous to those used in the proof of Proposition 6, it is immediate to show that the stochastic process induced in the present case is ergodic and thus long-run values are uniquely defined.

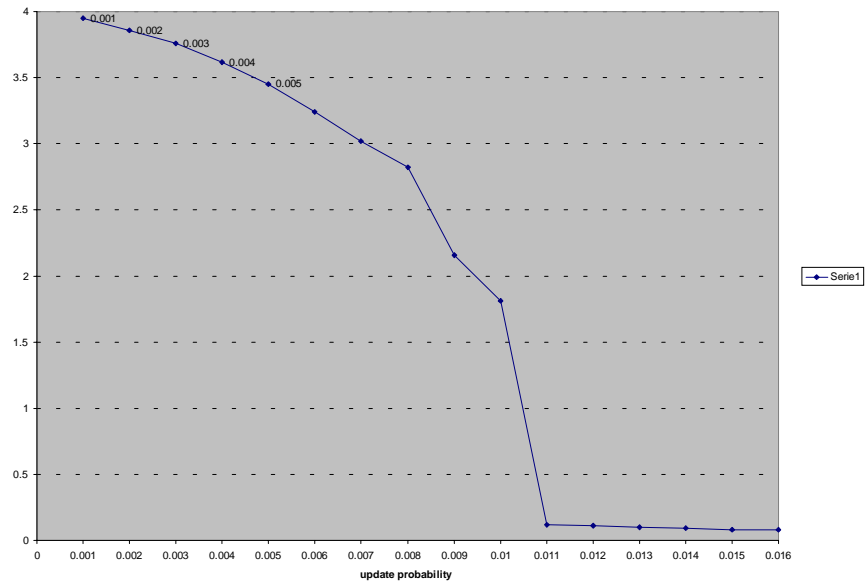


Figure 14: No network effects: payoff volatility and long-run average degree

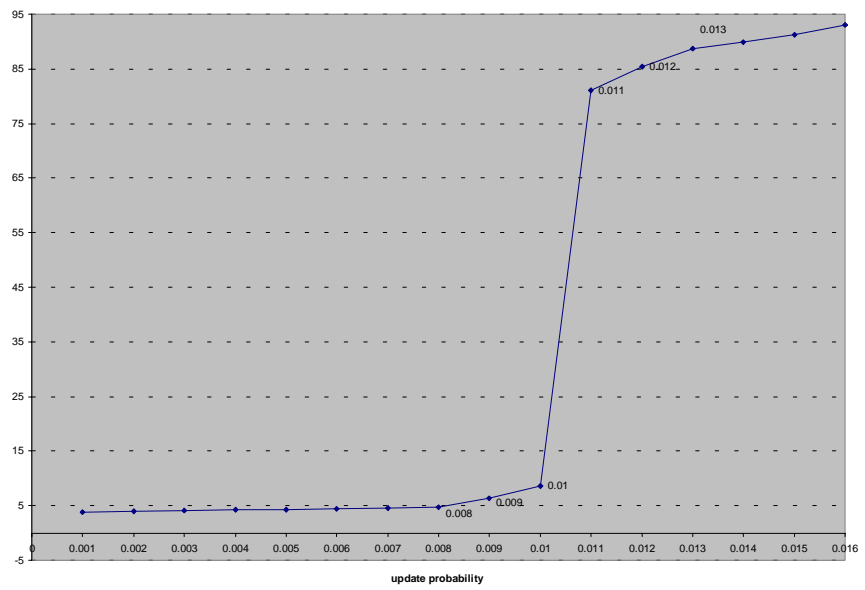


Figure 15: No network effects: payoff volatility and long-run average neighbor distance

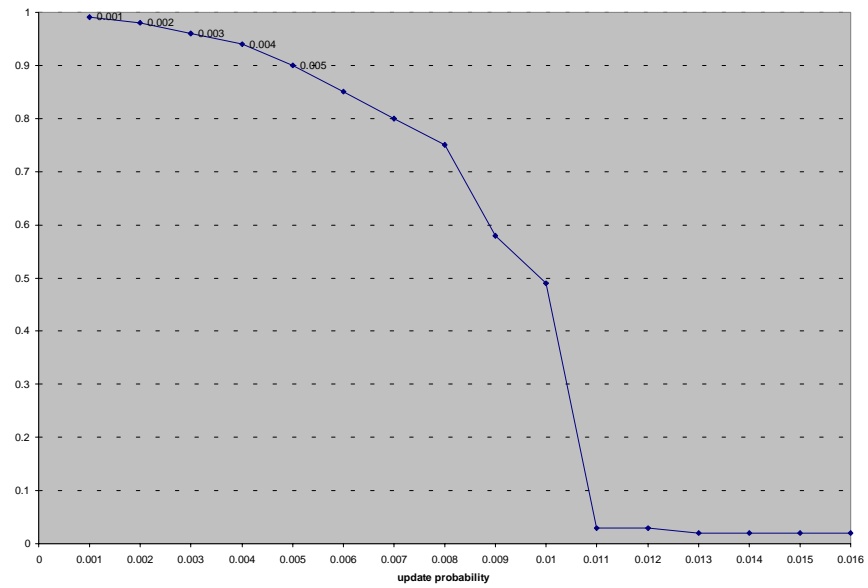


Figure 16: No network effects: payoff volatility and long-run share of largest component

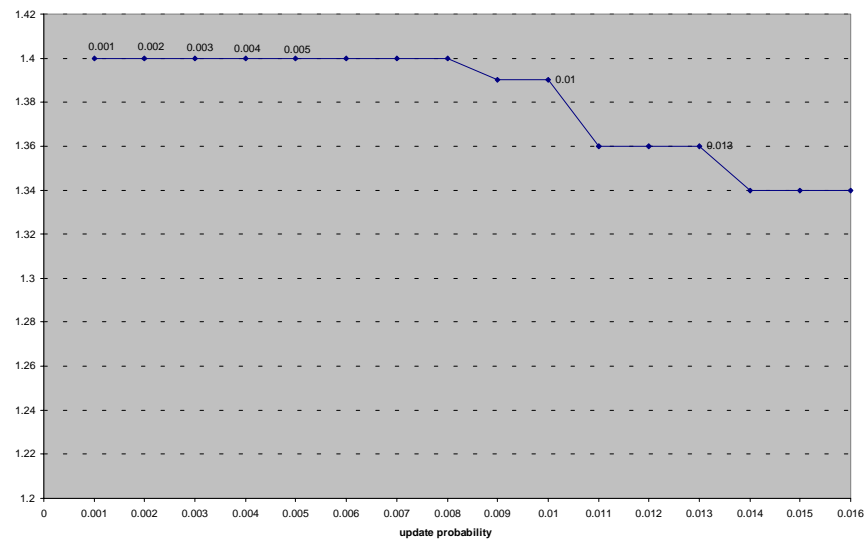


Figure 17: No network effects: payoff volatility and long-run average payoff

for ε that, again, is shorter than in the original framework by an order of magnitude. On the other hand, the “shape” of the curve tracing this change is roughly concave (i.e. with increasing differences), in contrast to the convex shape displayed by the original setup where further increases in ε led to progressively less significant changes in long-run magnitudes.

3. The behavior of average neighbor distance in this case also displays a marked contrast with the evidence obtained when players could rely on network strategic effects. Then, we emphasized that an increase in ε leads to rising network cohesiveness (i.e. decreasing neighbor distance), as the endogenous way in which the population ends up partially offsetting the increased volatility. Now, however, since cohesiveness brings about no relevant payoff consequences, an increase in ε induces a corresponding increase in neighbor distance. In essence, this transition is a mirror image of the parallel change experienced by the connectivity of the network, which naturally has the effect of increasing all network distances. (Recall the point made in this connection when discussing Figure 9.)

4.2.3 A less rewarding environment

Now, we focus on the effect of alternative payoff conditions on the behavior of the model. As it turns out, there are little surprises in this respect, and the essential points can be gleaned from just considering a scenario where the original payoff conditions are subject to a simple, say downward, shift in the payoff support. Specifically, it is supposed that this support continues to have a unit width but its upper end $\bar{\zeta}$ (and consequently its lower end $\underline{\zeta}$) are decreased by a mere 0.05, i.e. $[\underline{\zeta}, \bar{\zeta}] = [0.35, 1.35]$. This is still consistent with the hypotheses of Proposition 6 (thus ergodicity is guaranteed), but the ex-ante probability that a fresh draw exceeds the threshold $\frac{\nu(1-\delta)}{\delta} = \frac{4}{3}$ for bilateral-based cooperation now reduces to $\frac{1}{60}$.

Under these conditions, the effect of payoff volatility on long-run behavior is described in Figures 18-21. Comparing each of these figures with their counterparts obtained under the original payoff conditions, one finds the same qualitative behavior in terms of how the long-run values of each of the four variables of interest depend on the volatility rate. In the present case, however, the curves that reflect such dependence are seen to experience a global downward shift of significant magnitude across all four cases.

Concerning average degree, network span, or average payoff per link, the observed downward shift is simply the intuitive reflection of the fact that, as the payoff conditions deteriorate, so must happen as well with the performance of the social system. In the case of neighbor distance, on the other hand, the downward adjustment observed for every value of ε is again an interesting indication of how the network dynamics endogenously responds to the underlying exogenous circumstances. Here, in particular, it leads to an increase in the cohesiveness of the social network if, for any given volatility rate, the payoff environment becomes less rewarding on average.

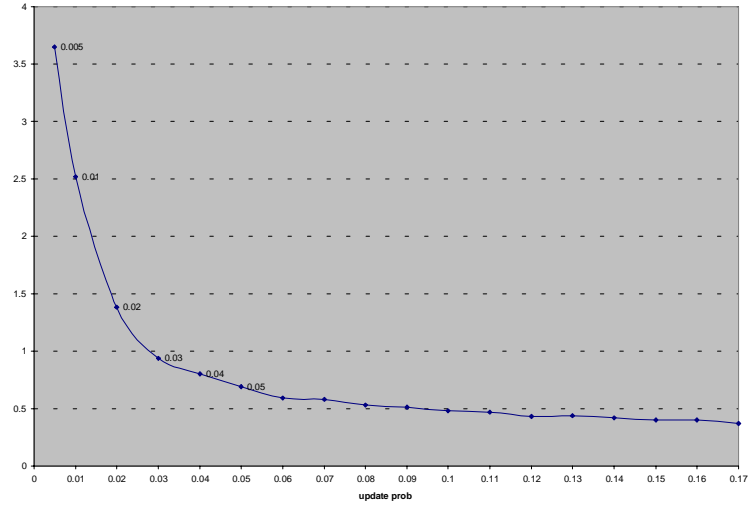


Figure 18: Low payoff support: volatility and long-run average degree

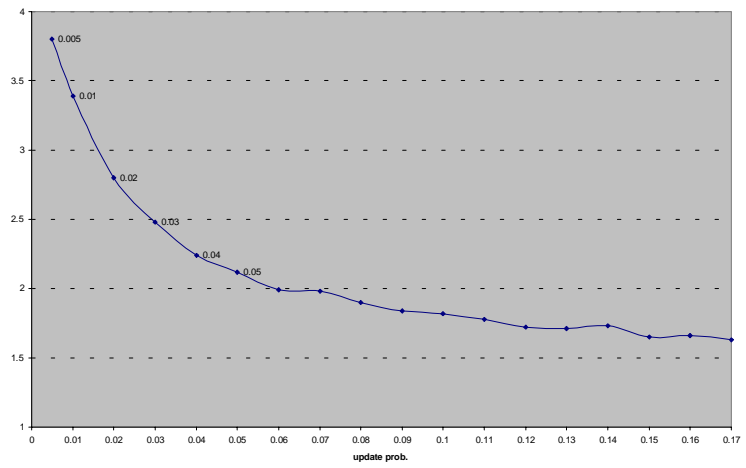


Figure 19: Low payoff support: volatility and long-run average neighbor distance

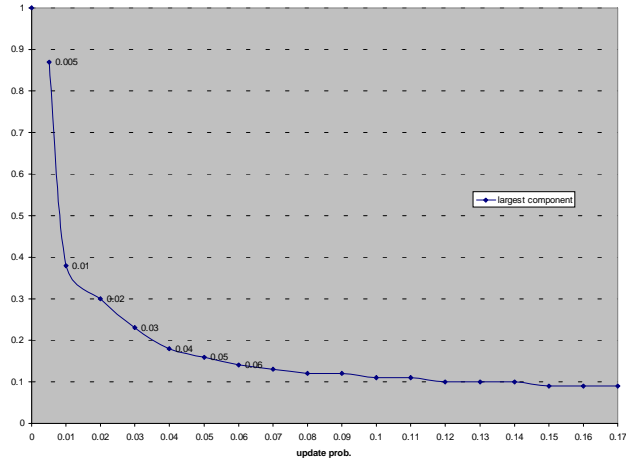


Figure 20: Low payoff support: volatility and long-run share of largest component

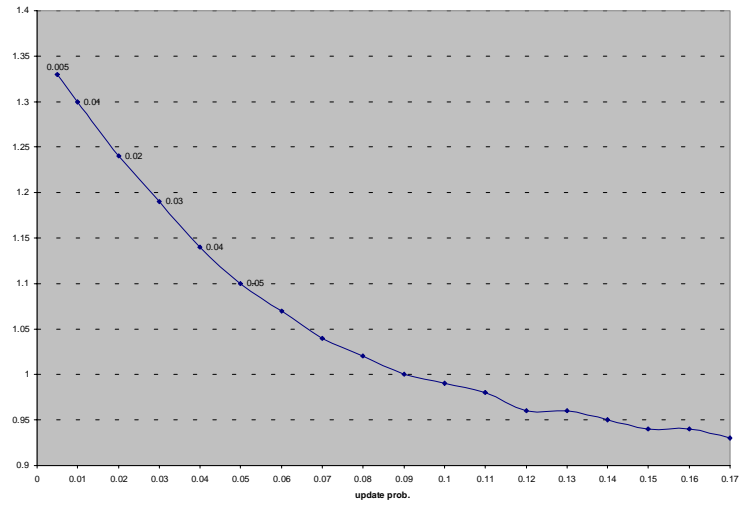


Figure 21: Low payoff support: volatility and long-run average payoff

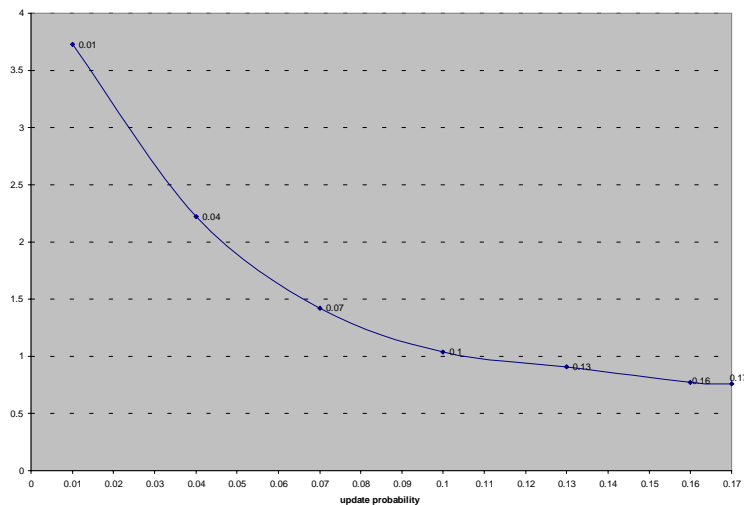


Figure 22: Search radius $\rho = 2$: volatility and long-run average degree

4.2.4 Short-radius search

One of the key features of the model has been to endow the social network with a crucial role concerning how the information diffuses throughout the population. This approach has been applied both to strategically-relevant information (i.e. past behavior of neighbors) as well as to fresh payoff opportunities. In the latter case, which arises when a player is at an “ordinary” (frequent) revision juncture, we have assumed that search is bound by her current component. In this respect, the interpretation has been that a revising agent can only be informed of payoff opportunities that arrive *through the network*, much as the information she receives pertaining to neighbors’ previous behavior. But, motivated by the idea that payoff-related knowledge may be complex and thus dissipate easily in bilateral communication, it is reasonable to contemplate the possibility that its “travelling distance” might be severely limited. Formally, this would amount to postulating that any fruitful search of new payoff opportunities is constrained by a certain radius, possibly of a short magnitude.

To explore the implications of this possibility, I have considered a variation of the original model where component-bound search (recall (2.a) in Subsection 2.2) only renders fresh payoff draws that are at a network distance of the revising agent lower than some pre-specified search radius ρ . For a starkest contrast, Figures 22-25 summarize the results for the lowest meaningful value of this radius, $\rho = 2$, and the parameter values used in the benchmark scenario.¹³

Interestingly, no major changes are found at a qualitative level, when the

¹³It can be readily verified that the argument used in Proposition 6 to establish ergodicity is independent of whether or not the component-bound search is restricted to a given radius.

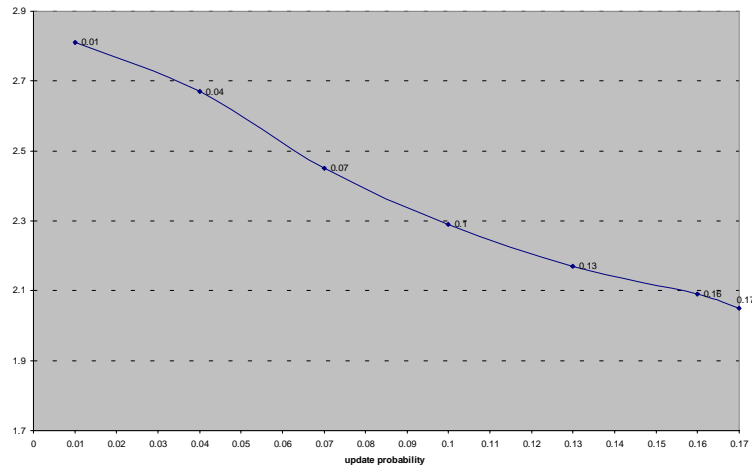


Figure 23: Search radius $\rho = 2$: volatility and long-run average neighbor distance

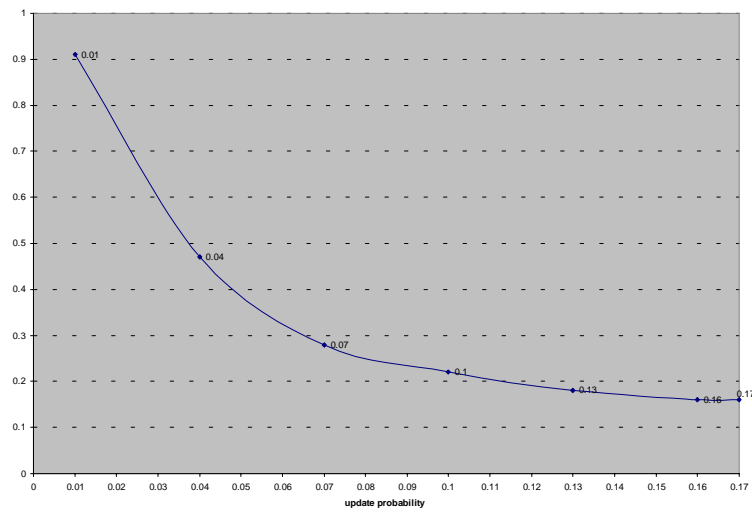


Figure 24: Search radius $\rho = 2$: volatility and long-run share of largest component

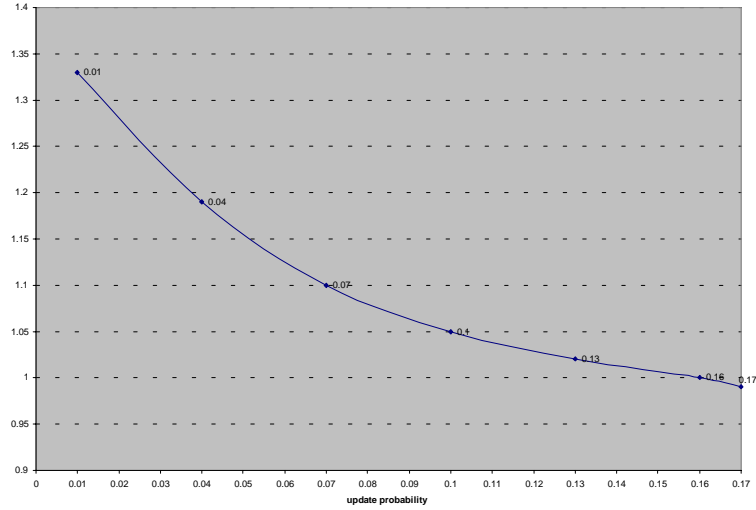


Figure 25: Search radius $\rho = 2$: volatility and long-run average payoff

present results under a short search radius are compared with those for the original model, as depicted in Figures 8-13. Focusing on the details, however, some intuitive differences do arise, some of which are worth noting. First, as one would expect, the short search radius translates itself into significantly lower neighbor distances, at all values of ε – i.e. a global downward shift of the corresponding curve (cf. Figure 23). Naturally, this larger cohesiveness in turn induces a larger connectivity (long-run average degree), which is a direct reflection of the stronger network effects that now impinge on players’ strategic considerations (cf. Figure 22). Finally, Figure 24 displays no significant effects on the largest-component shares observed at different volatility rates, while Figure 25 does show systematic negative differences with the average payoff earned per link in the benchmark scenario. The latter contrast represents an interesting complement to that observed with respect to network connectivity. While connectivity rises with a short search radius because of a stronger link stability (afforded, as explained above, by an increasing cohesiveness) the average payoff earned per existing link falls due to the narrower search imposed by that short radius.

5 Mean-field analysis

In this section, our objective is to undertake a so-called mean-field analysis of the model, akin to what is customarily done in statistical physics for large complex systems of interacting entities. In a nutshell, what underlies this approach is the presumption that, in the presence of a large number of stochastic inter-

actions, the aggregate (or average) behavior of the system can be reasonably well understood (or approximated) in expected terms – i.e. through a simpler *deterministic* representation of the process that embodies, on the one hand, an *average description* of the state of the system and defines, on the other hand, its law of motion in *expected (or mean-field)* terms.

As mentioned in the Introduction, such an analysis is carried out in more detail in the companion paper of Marsili, Slanina and Vega-Redondo (2002). Here, I simply outline part of it, the focus being to shed some new light on the conclusions derived from the numerical analysis of the model. More specifically, the main objective will be to study the dynamics of the network connectivity (i.e. the average degree) and its key dependence on the rate of payoff volatility. Concerning the latter, for example, we hope to gain some insight on why the effect of payoff volatility is affected, both quantitatively and qualitatively, by the presence or absence of strategic network effects. Naturally, much of what will be done with respect to network connectivity may be suitably extended to the study of other long-run variables, e.g. network span or average payoffs.

First, we ignore the components of the social dynamics that are concerned with payoff update and link creation through search. We start, therefore, by considering that part of the dynamics that deals with the gradual removal of pairwise-unstable links. Given any *possible* link $ij \in N \times N$, let $\lambda_{ij}(t) \in \{0, 1\}$ specify whether this link is in place at t (an event which is signified by $\lambda_{ij}(t) = 1$) or not (denoted by $\lambda_{ij}(t) = 0$). Then, for any such link ij , the dynamics of link-removal (due to pairwise instability) leads to the following law of motion:

$$\lambda_{ij}(t+1) = \lambda_{ij}(t) \cdot \mathcal{I}[\zeta_{ij}(t+1) - \beta_{ij}(t+1)], \quad (8)$$

where:

- $\zeta_{ij}(t+1)$ is the cooperation payoff prevailing at $t+1$ for the game played by i and j (it may be normalized to zero if there is no link connecting i and j),
- $\mathcal{I}[\cdot]$ is an indicator function defined by $\mathcal{I}[y] = 0$ if $y < 0$ and $\mathcal{I}[y] = 1$ if $y \geq 0$, and
- $\beta_{ij}(t+1) = \max\{\varphi_{ij}(t+1), \varphi_{ji}(t+1)\}$ with

$$\varphi_{ij}(t+1) = \frac{1-\delta}{\delta}\nu - \sum_{k \in N_i(t) \setminus \{j\}} \delta^{d^{i(k,j)}(t+1)} (\zeta_{ij}(t+1) + \frac{1-\delta}{\delta}\nu) \quad (9)$$

and $\varphi_{ji}(t+1)$ is defined reciprocally.

In the adjustment rule (8), the term $\beta_{ij}(t+1)$ quantifies the strength of the “strategic network effects” bearing on the link ij . Those effects are to be compared with $\zeta_{ij}(t+1)$ – i.e. must exceed this payoff – in order to confirm

the pairwise-stability of the link (and thus its remaining in place). To see this, simply note that the implied conditions for this state of affairs:

$$\begin{aligned}\zeta_{ij}(t+1) &\geq \varphi_{ij}(t+1) \\ \zeta_{ij}(t+1) &\geq \varphi_{ji}(t+1)\end{aligned}$$

are a mere rewriting of the pairwise-stability condition (5), as it pertains the link ij .

Next, we introduce a second component of the dynamics, namely, random payoff update. As formulated in Subsection 2.2, this process subjects every existing link to an independent probability ε of having its payoff updated afresh. Building upon the formulation used in (8), the expected law of motion that includes both payoff update and consequent link removal may be concisely written as follows:

$$\mathbf{E}_t[\lambda_{ij}(t+1)] = (1 - \varepsilon)\lambda_{ij}(t) + \varepsilon \lambda_{ij}(t) \Pr\{\zeta_{ij} \geq \beta_{ij}(t)\}, \quad (10)$$

where the implicit assumption here is that the expectation on β_{ij} is “static”, i.e. $\mathbf{E}_t[\beta_{ij}(t+1)] = \beta_{ij}(t)$. Therefore, every preceding link which is not subject to a payoff update remains in place and, for those which do experience such an update, their probability of staying pairwise-stable is:

$$\Pr\{\zeta_{ij} \geq \beta_{ij}(t)\} = \int_{\beta_{ij}(t)}^{\bar{\zeta}} f_{\zeta}(z) dz,$$

where f_{ζ} is the (continuous) density that governs every payoff draw.

Now, the key approximation step that is undertaken by the mean-field approach is to rely on the *link-specific* law of motion (10) to posit an average formulation resting on average quantities. Let $\kappa(t)$ stand for the average degree at t , i.e.

$$\kappa(t) = \frac{1}{n} \sum_{i \in N} \sum_{j \neq i} \lambda_{ij}(t).$$

Then, if we suppose that the network effects can be suitable averaged across players through a common $\hat{\beta}(t)$, the aggregate mean-field counterpart of (10) may be written as follows:

$$\begin{aligned}\mathbf{E}_t[\kappa(t+1)] &= \mathbf{E}_t \left[\frac{1}{n} \sum_{i \in N} \sum_{j \neq i} \lambda_{ij}(t+1) \right] \\ &= \mathbf{E}_t \left[\frac{1}{n} \sum_{i \in N} \sum_{j \neq i} \lambda_{ij}(t) \right] \left[(1 - \varepsilon) + \varepsilon \int_{\hat{\beta}(t)}^{\bar{\zeta}} f_{\zeta}(z) dz \right], \\ &= \left[(1 - \varepsilon) + \varepsilon \int_{\hat{\beta}(t)}^{\bar{\zeta}} f_{\zeta}(z) dz \right] \kappa(t),\end{aligned}$$

which may be rewritten as follows:

$$\begin{aligned}\mathbf{E}_t[\kappa(t+1)] &= \left[1 - \varepsilon + \varepsilon \left(1 - \int_{\underline{\zeta}}^{\hat{\beta}(t)} f_{\zeta}(z) \, dz \right) \right] \kappa(t) \\ &= [1 - \varepsilon \int_{\underline{\zeta}}^{\hat{\beta}(t)} f_{\zeta}(z) \, dz] \kappa(t).\end{aligned}$$

Finally, we complete the dynamics by adding the process of search and creation of new links. In the spirit of our mean-field approach, this dynamics of link creation is captured through a certain function $\phi(\kappa)$ of the current degree $\kappa \in [0, 2m]$, where recall that m is the maximum number of links that any given individual can support (therefore, $2m$ is the maximum average degree). For technical convenience, $\phi : [0, 2m] \rightarrow \mathbb{R}$ is assumed twice differentiable. Conceptually, this function is to be interpreted as embodying the rate of link creation that is associated to component-bound search, the “ordinary” (and more frequent) mechanism for creating new links. We postulate, therefore, that it satisfies the following boundary conditions:

$$\phi(0) = \phi(2m) = 0. \quad (11)$$

On the one hand, the condition that $\phi(\kappa)$ should vanish at $\kappa = 0$ is simply a reflection of the fact that, in an empty network, component-bound search yields no fresh opportunities. On the other hand, the additional condition that $\phi(\kappa)$ must vanish at $\kappa = 2m$ is a mere consequence of the limitation assumed on the number of active links that any single player can support – when every player is exhausting all her linking possibilities (thus the average degree is at its maximum), no further *net* increase is possible. To simplify the discussion, we shall also complement (11) with the following two additional conditions:

$$\phi'(0) = 0 \quad (12)$$

$$\phi(\kappa) > 0, \forall \kappa \in (0, 2m). \quad (13)$$

The first one is not crucial, but facilitates the discussion by making $\kappa = 0$ a clear-cut reference point in the analysis.¹⁴ The second condition, on the other hand, merely embodies the natural idea that as long as there is still some capacity for further connections, search always contributes, in expected terms, some positive “drift” of link creation.

Combining the three component of the dynamics considered (link removal, payoff update, and link creation), the mean-field law of motion for the network average degree can be compactly written as follows:

$$\kappa(t+1) = [1 - \varepsilon \int_{\underline{\zeta}}^{\hat{\beta}(t)} f_{\zeta}(z) \, dz] \kappa(t) + \phi(\kappa(t)), \quad (14)$$

¹⁴If this condition did not hold, a small value of κ (its magnitude associated to the frequency rate of search) would fulfill the same role as zero in the ensuing analysis.

where, to in order to obtain a deterministic system, we make the further simplification of substituting actual for expected motion in the LHS of the above expression. This leads to a one-dimensional autonomous system in $\kappa(t)$, which is the basis of our ensuing analysis.

A detailed study of the above dynamics requires positing some explicit functional dependence of $\hat{\beta}(t)$ on $\kappa(t)$. Of course, one obvious possibility in this respect would be to suppose that $\hat{\beta}$ is constant, i.e. independent of κ . This, indeed, is what characterizes the scenario where no strategic network effects are allowed, in which case the mean-field law of motion may be written as follows:

$$\kappa(t+1) = [1 - \varepsilon \int_{\underline{\zeta}}^{\nu \frac{(1-\delta)}{\delta}} f_{\zeta}(z) dz] \kappa(t) + \phi(\kappa(t)). \quad (15)$$

The above simple formulation is a direct consequence of the fact that, under no strategic network effects, a link can support cooperation if, and only if, it exceeds the threshold $\nu \frac{(1-\delta)}{\delta}$. But, in the strategic scenario of interest (i.e. that of the original model), the key feature is that $\hat{\beta}$ and κ are functionally related, a consideration that cannot be possibly ignored if the mean-field approach is to respect the essential modelling aspects of original framework.

Recall that, for each ij , $\beta_{ij}(t+1) = \max\{\varphi_{ij}(t+1), \varphi_{ji}(t+1)\}$. Therefore, in line with our present approach, it is natural to postulate that

$$\hat{\beta}(t) = A \langle \varphi_{ij}(t) \rangle_{ij \in g(t)}$$

for some $A \geq 1$, where $\langle \cdot \rangle$ stands for suitably computed averages. Then, in view of (9), expected magnitudes may be used to replace link-specific values to write:

$$\begin{aligned} \hat{\beta}(t) &= A \left\langle \frac{1-\delta}{\delta} \nu - \sum_{k \in N_i(t) \setminus \{j\}} \langle \delta^{d^\ell(u,v)(t)} \rangle_{\ell,u,v} \left[\langle \zeta_{uv}(t) \rangle_{uv \in g(t)} + \frac{1-\delta}{\delta} \nu \right] \right\rangle_{ij \in g(t)} \\ &= A \left\{ \frac{1-\delta}{\delta} \nu - \kappa(t) \langle \delta^{d^\ell(u,v)(t)} \rangle_{\ell,u,v} \left[\langle \zeta_{uv}(t) \rangle_{uv \in g(t)} + \frac{1-\delta}{\delta} \nu \right] \right\}. \end{aligned}$$

If, to be parsimonious, we replace the average magnitudes $\langle \delta^{d^\ell(u,v)(t)} \rangle_{\ell,u,v}$ and $\langle \zeta_{uv}(t) \rangle_{uv \in g(t)}$ by constant (i.e. time-invariant values), say $\hat{\delta}$ and $\hat{\zeta}$ respectively, we may write the above expression as follows:

$$\hat{\beta}(t) = A \left\{ U - \hat{\delta} \left[\hat{\zeta} + U \right] \kappa(t) \right\}, \quad (16)$$

where the notational shorthand $U \equiv \frac{1-\delta}{\delta} \nu$ is used.

For concreteness, let us adhere to the payoff conditions prevailing in the simulations and suppose that the cooperation payoffs ζ_{ij} are distributed uniformly on the support $[\underline{\zeta}, \bar{\zeta}]$ satisfying $\bar{\zeta} - \underline{\zeta} = 1$. Then, introducing (16) in (14), we arrive at the following expression:

$$\kappa(t+1) = \left\{ 1 - \varepsilon \cdot \max \left[0, A \left(U - \hat{\delta} \left(\hat{\zeta} + U \right) \kappa(t) \right) - \underline{\zeta} \right] \right\} \kappa(t) + \phi(\kappa(t)), \quad (17)$$

which is, under the contemplated assumptions, the mean-field approximation of the network-degree dynamics corresponding to the original model. As advanced, we are interested in comparing its induced dynamics with that resulting from a context where no strategic network effects are allowed. For this latter case, the suitable mean-field formulation is:

$$\kappa(t+1) = [1 - \varepsilon(U - \underline{\zeta})] \kappa(t) + \phi(\kappa(t)), \quad (18)$$

which is obtained from (15) by introducing the aforementioned uniformity assumption on the density f_{ζ} .

To start our discussion, an immediate conclusion that follows from mere inspection of (17) is that, at least for low values of κ , higher values of ε bring about a corresponding decrease network connectivity. That is, the higher is the volatility rate, the lower is (*ceteris paribus*, if the current level is not too high) the average degree prevailing in the next period. This represents a clear-cut manifestation of the detrimental effect of payoff volatility that, throughout our discussion, has played a central role in the motivation and analysis of the model.

Naturally, one expects that such negative bearing of volatility on the accumulation of social capital should have a more powerful effect at low levels of average connectivity, when network effects can play less of an offsetting role. This, indeed, is highlighted by the dynamics given by (17), where the point $\kappa = 0$ always turns out to be a locally stable equilibrium. (Here, of course, we make the assumption that $U \equiv \frac{1-\delta}{\delta} \nu > \underline{\zeta}$ and use that $\phi(0) = 0$.) Worded somewhat differently, this latter conclusion can be understood as suggesting that a *discrete* “upward shift” is always required if an *empty* configuration is to evolve into a social network with any *positive* connectivity. In the context afforded by our mean-field approximation, such an upward shift is to be conceived as the outcome of occasional “global” search, a component of the dynamics that, as explained, is kept only in the background.

Polar to the previous considerations, one would also expect that network configurations that display high levels of κ should prove quite resilient to volatility. And again, the mean-field dynamics (17) support this conjecture and also prove quite useful in understanding its underpinning. First, note that if κ may reach an arbitrarily large level (i.e. if m is sufficiently large), then the point $\kappa = 2m$ corresponding to maximally connected network is sure to be locally stable in terms of the mean-field dynamics. Specifically, the condition for this state of affairs to materialize is simply that

$$A \left(U - \hat{\delta} \left(\hat{\zeta} + U \right) 2m \right) - \underline{\zeta} < 0,$$

for which a straightforward sufficient condition is that

$$\hat{\delta} > \frac{1}{2m}. \quad (19)$$

Recall that $\hat{\delta}$ was identified with $\left\langle \delta^{d^{\ell}(u,v)} \right\rangle_{\ell,u,v}$, an average “effective discount rate” reflecting both the underlying discount rate δ and the time steps required

for information to move across two typical neighbors. Condition (19) does not appear to be demanding. For example, in our benchmark simulations, if we were approximate $\hat{\delta}$ by the number that results from raising the discount rate δ ($= 3/4$) to the long-run average neighbor distance (which is never higher than 4, even for very low values of ε), a lower bound for it would be 0.3. This rather loose lower bound, in turn, is greater than $\frac{1}{2m} = \frac{1}{4}$, which suggests that the conditions displayed by the benchmark scenario are consistent with condition (19).

Thus assume that this condition holds. Then, the dynamics (17) has the two extreme points, $\kappa = 0, 2m$, each defining locally stable (i.e. “dynamically robust”) configurations. Under what conditions are these points the only robust equilibria? To answer this question, rewrite (17) as follows:

$$\kappa(t+1) - \kappa(t) = -\varepsilon \cdot \max \left[0, A \left(U - \hat{\delta} \left(\hat{\zeta} + U \right) \kappa(t) \right) - \underline{\zeta} \right] \kappa(t) + \phi(\kappa(t)), \quad (20)$$

and have the first term of its RHS define the function $h(\kappa) \equiv -\varepsilon \cdot \max \left[0, A \left(U - \hat{\delta} \left(\hat{\zeta} + U \right) \kappa \right) - \underline{\zeta} \right] \kappa$. When this latter function is strictly negative (and therefore the positive truncation does not apply), it is strictly convex. Thus, if the other function in the RHS of (20), $\phi(\kappa)$, does not display an “excessive opposite curvature” which may offset that convexity for low κ (see Proposition 9 for a suitable bound), at most one *interior* equilibrium of exists and the question posed above can be answered in the affirmative. That is, the two end points of the state space are the sole robust (asymptotically stable) states in this case. Furthermore, it then follows that the state space can be essentially partitioned into two disjoint basins of attraction, K_0 and K_{2m} , which are separated by an *interior* threshold $\tilde{\kappa}$. When the process starts below this threshold (i.e. $\kappa(0) \in K_0$), it converges to the lower end of the state space given by $\kappa = 0$. Instead, when the initial condition $\kappa(0)$ lies above the threshold (i.e. belongs to K_{2m}), the induced path converges to the upper end $\kappa = 2m$.

To sum up the former discussion, it is useful to gather formally its main points in the following result, whose complete proof is omitted for the sake of brevity.

Proposition 9 *Consider the discrete-time dynamics (17) defined on $[0, 2m]$ with $U > \underline{\zeta}$ and $\hat{\delta} > \frac{1}{2m}$. Further assume that $\phi(\cdot)$ satisfies (11)-(13) and $\phi''(\kappa) > -\varepsilon \hat{\delta} \left(\hat{\zeta} + U \right)$ for all $\kappa \leq \frac{U}{\hat{\delta}(\hat{\zeta}+U)}$.¹⁵ Then, there is a certain threshold value $\tilde{\kappa} \in (0, 2m)$ such that:¹⁶*

- (a) if the initial point $\kappa(0) < \tilde{\kappa}$, then $\lim_{t \rightarrow \infty} \kappa(t) = 0$;
- (b) if the initial point $\kappa(0) > \tilde{\kappa}$, then $\lim_{t \rightarrow \infty} \kappa(t) = 2m$.

¹⁵In fact, it turns out that the lower bound on $\phi''(\kappa)$ can be required only for $\kappa \leq \frac{AU - \underline{\zeta}}{\hat{\delta}A(\hat{\zeta} + U)} (< \frac{U}{\hat{\delta}(\hat{\zeta} + U)})$.

¹⁶Note, of course, that (a) and (b) make, respectively, the points $\kappa = 0$ and $\kappa = 2m$ be locally (i.e. asymptotically) stable.

The above result provide an approximate analysis of the network dynamics (in particular, its degree) through the expected law of motion of a typical node/player embodying average considerations. The main implications of this analysis is that there are just two polar configurations which are possible attractors of the model’s “representative behavior”. On the one hand, if the starting connectivity conditions are dense enough, the adjustment dynamics leads to a situation of maximum connectivity. Instead, if the initial average density does not reach the required threshold, the mean-field dynamics leads towards the eventual vanishing of all links.

Of course, the “representative-node” fiction that is used to motivate the mean-field barouche is a heroic construct, while the average considerations that are supposed to guide its behavior can only represent a compact way of identifying the essential dynamic forces governing a very complex system. The insights thus obtained, however, may be viewed as an expression of the tendencies that govern the dynamics of the different network components. (Recall that the postulated mean-field dynamics only embodies component-based forces.) And, in this vein, a first insight that arises is that, as a general tendency, one should expect to find the network divided into two kinds of components: those where the average connectivity reaches its maximum level and others displaying no links whatsoever (i.e. singleton components). In a sense, therefore, the prediction is that the positive feedback effects that underlie the network dynamics should either bring about the full connecting potential or, instead, must prove completely unable to trigger a successful buildup. Indeed, this is the behavior observed in the simulations, as illustrated, for example, in Figures 10 and 11, or suggested by a comparison of the figures which depict the average degree and largest-component share (e.g. Figures 8 and 12).¹⁷

But then, building upon this line of thought, a complementary prediction is that the relative sizes of the basins of attraction for the two polar configurations ($\kappa = 0, 2m$) should be in line with the relative fraction of nodes observed in each of the two types of components – i.e. maximally connected or empty ones. This in turn suggests studying how the threshold $\tilde{\kappa}$ which marks those relative sizes depends on the volatility rate ε , a dependence that should help clarify the contrast between the different qualitative behavior observed in the presence or absence of strategic network effects. In this respect, it is straightforward to rely on the analysis that led to Proposition 9 to arrive at the following result.

Proposition 10 *Consider the discrete-time dynamics (17) under the same assumptions as in Proposition 9 and denote by $\tilde{\kappa}(\varepsilon)$ the threshold established in it, as a function of the underlying value of ε . Then, for any $\varepsilon' > \varepsilon$ (keeping the rest of the parameters fixed), $\tilde{\kappa}(\varepsilon')$ changes continuously with ε' in an increasing fashion.*

¹⁷For example, if one computes the ratio between the average degree and the fraction of nodes involved in non-singleton components, the result is always barely below 4 (which is equal to the maximum average degree in this case, where $m = 2$).

The previous result provides a formal basis to understand the negative dependence of the network connectivity on the underlying rate of payoff update. This dependence, in turn, is what explains the negative effect of payoff volatility on other measures of performance such as, for example, the average payoff earned across existing links. An important feature of this relationship between the volatility rate and average degree is that it is gradual – or, expressed in mathematical terms, continuous. This contrasts with the behavior observed in the alternative scenario where players’ strategic interaction relied on no network effects. In this case, the dependence on ε was still negative and gradual within a certain low range but, at a certain point, an abrupt change in long-run performance occurred associated to a small change in this parameter. Thus, in comparison with the situation observed for the benchmark scenario, this latter evidence could be understood as the indication of a discontinuous regime change caused by a slight change in ε .¹⁸

To shed light on this contrast, let us return to the mean-field description of the scenario with no strategic network effects (cf. (18)), which is convenient to rewrite as follows:

$$\kappa(t+1) - \kappa(t) = \phi(\kappa(t)) - \varepsilon(U - \zeta) \kappa(t). \quad (21)$$

To draw a clearer comparison with the benchmark scenario, it is useful to introduce some new notation. Given any volatility rate $\varepsilon > 0$, denote by $\Omega(\varepsilon)$ the limit points which have a basin of attraction of positive measure in the benchmark scenario, and let $\tilde{\Omega}(\varepsilon)$ be its counterpart for the scenario with no strategic network effects. From Proposition 9, we know that, under the maintained assumptions, we have:

$$\Omega(\varepsilon) = \{0, 2m\}$$

for all $\varepsilon > 0$.

Turning now to the alternative dynamics given by (21), it is clear (again under the assumptions of Proposition 9) that, if ε is low enough albeit positive, $\tilde{\Omega}(\varepsilon)$ includes not only $\kappa = 0$ but also some other positive values of κ . Denote by

$$\tilde{\kappa}(\varepsilon) \equiv \max \{ \kappa : \kappa \in \tilde{\Omega}(\varepsilon) \}$$

the largest limit point for any given $\varepsilon > 0$. Then, for low ε , is close to $2m$ and changes continuously with ε . Now let ε^* be the maximum value of ε for which $\tilde{\kappa}(\varepsilon) > 0$. Precisely, ε^* is the slope of the smallest minimum cone rooted at the origin which encompasses the region in \mathbb{R}_+^2 under the function ϕ . Under our maintained assumptions, it is straightforward to see that $\varepsilon^* < 1$. And, by construction, it follows that, for all $\varepsilon > \varepsilon^*$, $\tilde{\kappa}(\varepsilon) = 0$ (or, equivalently, $\tilde{\Omega}(\varepsilon) = \{0\}$), which implies a discontinuous change in the limit behavior of the dynamics (21). This discontinuity in the limit behavior of the mean-field

¹⁸In statistical physics, this behavior is often encountered and usually referred to as a “phase transition.”

dynamics should in turn be mirrored by a corresponding abrupt change in the long-run behavior of the stochastic dynamics of the full-scale system.

We find, therefore, that the qualitative contrast found in the simulations between the ε -dependence displayed by the social dynamics with, or without, strategic network effects is reconfirmed by our present analytical approach. In essence, what we obtain is theoretical support for the intuitive idea that network effects may be an effective way for a society to respond effectively to a volatile environment.

6 Related literature

In this section, I schematically review the relationship of the present paper to a variety of different literatures.

The approach pursued here is similar in spirit to that of the evolutionary literature, where players are assumed to interact through a certain game and the long-run configuration is obtained through a gradual stochastic process of learning and adjustment. The seminal papers in this line of research are Kandori, Mailath, and Rob (1993) and Young (1993). In these papers, the interaction is global – all agents play with every other player in the population, or at least have the same probability of doing so. This feature was altered in Ellison (1983), where each agent’s interaction is local, play restricted to immediate neighbors on a *fixed* regular network (or grid). More recent work by Droste *et al.* (1999), Goyal and Vega-Redondo (1999), or Jackson and Watts (1999) have extended the analysis to a context where the network is not fixed but co-evolves with players’ game decisions. This latter work is, therefore, the closest in motivation to the present endeavour.

The aforementioned evolutionary literature has typically considered contexts where the game played by the population is a simple bilateral game in strategic form – often a coordination game. Players are required to play the same action in each interaction, since otherwise the network would have no influence on strategic behavior. This contrasts with our present framework where the decisions are adopted independently for each of the games being played and the network influence on players’ behavior is only indirect through the way in which strategically-relevant information diffuses through the population.

There is, moreover, the booming body of literature whose specific concern is the study of pure models of network formation, in contexts where players do not have a separate strategic decision in terms of constituent game. One of the earliest papers in this field was Aumann and Myerson (1989), with the more recent paper by Jackson and Wolinsky (1996) having played an important role in reviving interest in this topic. Whereas the approach of these papers is mostly static,¹⁹ Bala and Goyal (2000) represents an explicit dynamic approach to the problem, and is thus more in line with the present efforts.

¹⁹For example, Jackson and Wolinsky rely on a notion of pairwise-stability that bears a close relationship to that introduced in Definition 1.

In the vast area of repeated games, there are two papers, Kandori (1992) and Ellison (1994), which share some motivation with our approach. They propose a model where a large population of players are repeatedly and randomly matched to play a Prisoner Dilemma game. They find that, in this context, it may be still possible to induce cooperation through a social norm (equilibrium) that reacts to any deviation by punishing subsequent partners. In a sense, the relationship of these papers to our work is parallel to that displayed by the early evolutionary literature with fixed and global interaction structure: they embed players' interaction in a population context but abstract from the effect of social structure by postulating a fixed and global pattern of play.

Still in the area of repeated games, two additional related papers are Haag and Lagunoff (2000) and the already mentioned by Bernheim and Whinston (1990). The latter studies a model of multimarket collusion where a group of firms participating in some common set of markets may decide to link their behavior in any one of them to what has been observed in other markets. Naturally, this enhances the collusion (i.e. cooperation) potential, in ways analogous to those considered here. The key difference is that the flow of information is instantaneous and therefore there is no interesting notion of social network that may have a bearing on the situation. Instead, the paper by Haag and Lagunoff does study a setup where players are involved in repeated interaction with partners specified by some *given* social network. Its approach, however, is mostly normative and static, with an additional important difference being that players are forced to play the same action – cooperate or defect – with everyone of their neighbors. Therefore, the strategic implications of the (exogenously specified) network are directly given by the formulation of the population game (which requires a common action in all encounters) rather than being of an informational nature on a set of “decoupled” bilateral interactions. In a this light, our approach may be viewed as embedding the Kandori's and Ellison's aforementioned approach with independent bilateral interaction into a endogenously changing social network.

The study of social networks has hardly been a preserve of economists or game theorists. Rather, it has long been a primary object of study by sociologists or applied psychologists. Besides the research it has spawned in connection to the notion of social capital (which is discussed later), prominent sociologists such as Mark Granovetter (1973) or Ronald Burt (1982) have placed it at the center of sociological inquiry. For example, the notions of “weak ties” highlighted by Granovetter, or “structural holes” due Burt have given rise to a large body of theoretical and empirical work in sociology, which still continues to thrive. Other early and well-known research was carried out by Milgram (1967), an applied psychologist, who demonstrated through clever simple experiments the surprising low number of steps which tend to separate any two arbitrarily chosen individuals in many large social networks. Since then, most such networks are known to be of a so-called “small world” kind, a phenomenon that has recently attract much attention by physicists and other researchers interested in the study of complex systems. A strong stimulus to this literature was provided by Watts and Strogatz (1998), which triggered a booming and wide-ranging collection of

multidisciplinary work, in which still physicists such as Albert-Lazslo Barabasi (see Barabasi and Albert (1999)) or Mark Newman (see Newman *et al.* (2000)) play a leading role.

Finally, we close this brief review by referring to the literature and ideas that were amply invoked in the Introduction to motivate our model, namely, those connected to the notion of social capital. Rather than attempting a necessarily superficial survey of the vast and diverse range of research that goes under this heading, it should be more useful to focus on the work of James Coleman, arguably the “father” of this concept. He is also the author who appears to have conceived social capital more in line with the view espoused here – see Coleman (1988) or (1990, Ch. 12). For him, social capital is an inherently relational concept, best regarded as an attribute of the social network. It is the key factor explaining the intensity and stability of socio-economic interaction and also represents the basis of trust in repeated interaction (in particular, Coleman often uses the PD as the paradigmatic example). However, for such a trust to emerge, what he calls the closure of the social network (what we have called its cohesiveness) is generally key. He argues, moreover, that social capital is often underprovided, since the strong externalities associated to it are typically not internalized by individuals’ own link-investment decisions. Finally, he stresses that social capital is a stock which, left to itself, depreciates with time and that, if it is to be (re)built successfully, must have inter-agent relations enjoy a sufficiently stable environment. The reader will recognize in these points many of the features (both modelling and motivational) that have informed our efforts in this paper.

7 Summary and possible extensions

This paper has studied a stylized model of network formation in which players are involved in an infinitely repeated Prisoner’s Dilemma with each of her neighbors. Information on past behavior flows gradually along the network, a feature that impinges crucially on the range of network configurations that can be supported in a pairwise stable fashion. The underlying payoff conditions change over time, which in turn may affect the stability of established links and create the opportunity to form new ones. Our analysis has focused on the interplay between the emerging characteristics of the endogenous network, the long-run performance of the system, and the key parameters of the model – most importantly, the volatility rate at which the payoffs of current links is updated.

The main insights obtained can be summarized as follows. Payoff volatility cannot be too intense if a successfully build of the network is to be sustained in the long run. As the environmental conditions worsen (e.g. payoffs becomes more volatile or less rewarding), the network endogenously evolves in order to offset this turn of events by becoming more cohesive. At a certain point, however, even this mechanism proves unable to maintain other than a quite sparse network. This transition, which takes place both if players can rely on strategic network effects or not in supporting cooperation, appears much earlier

and is more abrupt in the latter case than in the former. These conclusions have been obtained through a “numerical computation” of the ergodic variables of the process, but further insight has been gained on them by a mean-field approximation of the basic model.

The model is to be conceived as a first step along an as-yet little trodden path. Thus let me finish with some suggestions as to some extensions and issues that are left for future research.

A natural extension would involve enriching the set of games under consideration, possibly to other sorts of simple bilateral games (e.g. coordination games) or playing the field contexts. Along these lines, one further possibility would be to suppose that, as found in the literature (see above), players are forced to take the same action in all games. The interplay between the network considerations brought about by this modification and the informational aspects considered here may yield novel insights.

Concerning payoff volatility, it would be interesting to allow for the (arguably realistic) possibility that the realizations induced by any fresh payoff update may be correlated in some dimension. For example, the payoffs pertaining a particular individual might display positive correlation (the reflection of a generally creative or well positioned player at some point in time) or, in a somewhat similar vein, new payoff draws could include an aggregate component. Any of these modifications is bound to yield important implications on the network dynamics and its long-run architecture.

Payoffs, on the other hand, might be subject to some exogenous or endogenous trend. Concerning the first possibility, it would be interesting to postulate that existing payoffs are affected by some negative drift, a phenomenon that could reflect an outside forward-moving environment that sets the benchmark for the payoffs obtained. As for a trend of an endogenous kind, it would be interesting to suppose that the payoffs earned are to be normalized by population-average payoffs or that, new payoffs are drawn according to a moving distribution anchored to average or frontier conditions. In either case, a supplementary competitive pressure would be added to the model that may well introduce new considerations, as well as endow the model with a genuinely dynamic (or growth) perspective.

To conclude, a final suggestion concerns the possibility of removing the fixed upper bound on the number of links that an agent can support but instead contemplate a cost to be incurred by the players involved (perhaps equally shared, or possibly with a larger cost share paid by the initiating party). This would make the number of links that is effectively supported an endogenous variable, and would quite naturally link the density of the network to prevailing payoff conditions.

References

- [1] Abreu, D. (1998): “On the theory of infinitely repeated games with discounting”, *Econometrica* **56**, 383-96.
- [2] Aumann R. and R. Myerson (1989): “Endogeneous formation of links between players and coalitions: an application of the Shapley value”, in A. Roth (eds) *The Shapley Value*. Cambridge University Press.
- [3] Bala, V. and S. Goyal (2000): “A non-cooperative model of network formation”, *Econometrica* **68**, 5, 1181-1231.
- [4] Barabasi, A.-L. and R. Albert (1999): “Emergence of scaling in random networks”, *Science* **286**, 509-12.
- [5] Bernheim, D. and M. Whinston (1990): “Multimarket contact and collusive behavior”, *Rand Journal of Economics* **21**, 1-26.
- [6] Burt, R.S. (1982): *Towards a Structural Theory of Social Action*, New York: Academic Press.
- [7] Coleman, J. S. (1988): “Social capital in the creation of human capital”, *American Journal of Sociology* **94**, S95-S120.
- [8] Coleman, J. S. (1990): *Foundations of Social Theory*, Cambridge, Ma: Harvard University Press.
- [9] Droste, E, R. Gilles, and K. Johnson (1999): “Endogenous interaction and the evolution of conventions”, mimeo, Tilburg University.
- [10] Durlauf, S. N. (1999): “The case ‘against’ social capital”, mimeo, University of Wisconsin.
- [11] Ellison, G. (1993): “Learning, local interaction, and coordination”, *Econometrica* **61**, 1047–1071.
- [12] Ellison, G. (1994): “Cooperation in the Prisoner’s Dilemma with anonymous matching”, *Review of Economic Studies* **61**, 567-88.
- [13] Goyal S. & F. Vega-Redondo (2002): “Network formation and social coordination”, mimeo, Erasmus University and Universidad de Alicante.
- [14] Granovetter, M.S. (1973): “The strength of weak ties”, *American Journal of Sociology* **78**, 1360-80.
- [15] Haag, M. & R. Lagunoff (2000): “Social norms, local interaction, and neighborhood planning”, mimeo, Department of Economics, Georgetown University.
- [16] Jackson, M. and A. Watts (2002): “On the formation of interaction networks in social coordination games”, *Games and Economic Behavior*, forthcoming.

- [17] Jackson, M. and A. Wolinsky (1996): "A Strategic model of economic and social networks", *Journal of Economic Theory* **71**, 44-74.
- [18] Kandori, M. (1992): "Social norms and community enforcement", *Review of Economic Studies* **59**, 63-80.
- [19] Kandori, M. and G. J. Mailath and R. Rob (1993): "Learning, mutation, and long run equilibria in games", *Econometrica* **61**, 29-56.
- [20] Lucas, R. (1988): "Making a miracle", *Econometrica* **61**, 251-72.
- [21] Marsili M., F. Slanina & F. Vega-Redondo (2002): in preparation.
- [22] Milgram, S. (1967): "The small-world problem", *Psychology Today* **2**, 60-67.
- [23] Newman, M.E.J., C. Moore, and D.J. Watts (2000): "Mean-field solutions of the small-world network model", *Physical Review Letters* **84**, 3201-04.
- [24] Watts D.J. and S. H. Strogatz (1998): "Collective dynamics of 'small-world' networks", *Nature* **393**, 440-42.
- [25] Woolcock, M. (2000): *Using Social Capital: Getting the Social Relations Right in the Theory and Practice of Economic Development*, Princeton, NJ: Princeton University Press.
- [26] Young, H. P. (1993): "The Evolution of Conventions", *Econometrica* **61**, 57-84.