Chapter 8
Homophily, Conformity, and Noise in the (Co-)Evolution of Complex Social Networks

George Ehrhardt, Matteo Marsili, and Fernando Vega-Redondo

8.1 Introduction

Social networks constitute the backbone underlying much of the interaction conducted in socioeconomic environments.[1] Therefore, when this interaction attains a *global* reach it must have, as its counterpart, the emergence of a social network with a wide range of overall (typically indirect) connectivity. Naturally, for such a social network to emerge, agents must be able to link profitably. But this in turn demands that they display similar – at least compatible – behaviour. Thus, for example, they must use coherent communication procedures, share key social conventions, or have similar technical abilities. Here, we may quote the influential work of Castells (1996).

> Networks are open structures, able to expand without limits, integrating new nodes as long as they are able to communicate within the network, namely as long as they share the same communication codes (for example, values or performance goals).

Reciprocally, of course, such a convergence of behaviour is facilitated by the range of social interaction being global rather than local or fragmented. This suggests the idea that the buildup of a global social network might be understood as the outcome of twin cross-reinforcing processes: one that facilitates the convergence of norms and behaviour, and another that extends the range of

F. Vega-Redondo (✉)
Department of Economics, European University Institute, Florence 50133, Italy

[1]Some important examples include labor markets (Granovetter 1974; Montgomery 1991; or Calvó-Armengol and Jackson 2004), trade arrangements (Kirman et al. 2000; Kranton and Minehart 2001), or networks of interfirm collaboration in either research and development (Hagedoorn 2002; Powell et al. 2005).

A. Reggiani and P. Nijkamp (eds.), *Complexity and Spatial Networks*, Advances in Spatial Science,
DOI: 10.1007/978-3-642-01554-0_8,

social connectivity. This mutual reinforcement also suggests that if such a global transition indeed takes place, it should be relatively *fast* and *resilient*.

To explore these ideas, Ehrhardt et al. (2006) – thereafter referred to as EMV – proposed a stylized model in which agents are involved in a local coordination game with their neighbours in a coevolving network. Two features characterize the dynamics. First, we postulated that while links vanish at a constant exogenous rate, new links are created only between agents who randomly meet and happen to be "coordinated", that is, display the same action/strategy in the coordination game. On the other hand, we assumed that, on the same time scale (that is, at a comparable rate), agents adjust their action towards that which maximizes the extent of coordination with their current neighbours. For brevity, we referred to the first feature as *homophily* and the second as *conformity*.

The way in which EMV conceive both homophily and conformity is particularly stark but also restrictive. In particular, both dynamic forces are implemented through noiseless mechanisms, which implies that, in the long run, the only links existing in the network are those connecting agents displaying the same action – that is, network components are homogeneous. This seems too extreme a setup, which make one wonder whether the analysis is robust to the introduction of some noise. Our aim here will be to conduct such a "robustness check" by studying a model where some persistent noise may perturb both the establishment of links and the adjustment of actions.

In a nutshell, our conclusion is that neither of these generalizations affects the main predictions of the model. As in the original benchmark model, we continue to observe:

1. Sharp qualitative transitions in network connectivity and coordination, as a discontinuous (upward and downward) response to slight changes in the underlying parameters beyond corresponding thresholds.
2. The transitions mentioned in (1) display hysteresis, that is, they are locally irreversible in the long run, even if the environmental parameters revert to their original values.
3. As a consequence of (2), there is a sizable range of parameter values for which the system exhibits long-run multiplicity, which is resolved depending on history or/and the initial conditions.

EMV observed the phenomenology (1)–(3) as the rate of volatility, the only significant parameter of the model, was varied gradually along its full range. We now confirm that, in the generalized model, the same qualitative behaviour concerns changes in the noise parameters now introduced, both concerning homophily and conformity. In fact, since their implications are fully parallel, we shall focus our present discussion on the parameter modulating the noise of action adjustment, as will be explained in detail below.

The remainder of the chapter is organized as follows. Section 8.2 presents the model, Sect. 8.3 carries out the analysis and motivates it, while Sect. 8.4 concludes with a recapitulation.

8.2 The Model

Let there be a certain population of agents, $P = \{1, 2, \ldots, N\}$, who interact bilaterally over time as specified by the evolving social network. Time is modelled continuously, with $t \in 0, \infty)$. At any t, the state of the system $\omega(t)$ consists of two items: (1) the social network *g(t)* that specifies the set of undirected links *ij* (=*ji*) prevailing at *t*; (2) the action profile $\alpha(t) \in A^M$, where $A = \{a_1, a_2, \ldots, a_q\}$ is the set of q possible actions.

Players adjust both actions and links over time. The dynamics is described by a continuous Markov process for the state $\omega(t)$, and is therefore completely determined by the rates governing all possible transitions $\omega \to \omega^{'}$. These transitions pertain to adjustments that involve (1) link creation, (2) link destruction, (3) action revision. We now describe each of these in turn.

Link creation: We posit that at a certain positive rate η each agent i receives a link creation opportunity. When such an opportunity arrives at some t, another agent j is randomly chosen in the population (all with the same probability). When no link exists between i and j (that is, $ij \notin g(t)$), the link ij is formed with probability one if

$$\alpha_i(t) = \alpha_j(t).$$

Otherwise, the link is formed with probability ε, conceived small.

Link destruction: It is assumed that existing links decay at a rate λ. This component of the process may be provided with different (non-exclusive) interpretations. For example, it may be conceived as a reflection of unmodelled environmental *volatility* that affects the value or feasibility – and thus the persistence – of some of the existing links.

Action revision: At every t, each agent i independently receives at a rate v the opportunity to revise her current action. If this revision opportunity materializes, then she chooses every possible action $a_r \in A$ with probability

$$\Pr(\alpha_i(t) = a_r) = \frac{1}{H} \exp\left[\beta \left| \left\{ \alpha_j(t) = a_r : ij \in g(t) \right\} \right| \right], \tag{8.1}$$

where

$$H \equiv \sum\nolimits_{r'=1}^{q} \exp\left[\beta \left| \left\{ \alpha_j(t) = a_{r'} : ij \in g(t) \right\} \right| \right]$$

is a normalization factor and $\beta \geq 0$ is a parameter that modulates the sensitivity of agents' adjustment to conforming (or coordinating) with the local environment. It can be understood as embodying a desire of the agents to play optimally in a local coordination game, where the *instantaneous* payoffs that may be obtained from each action are linear in the number of neighbours *currently* displaying that same choice.

The above formulation yields the model studied in EMV as a particular case when

$$\varepsilon = 1/\beta = 0,$$

that is, when the noise associated to link creation and action revision are both zero. As advanced, since the two sources of noise lead to totally analogous implications, we abstract from the former by still making $\varepsilon = 0$ while we focus our attention on the latter by assuming that β is finite. Note that, in this case, (8.1) becomes the specific exponential form that has been amply used in modern evolutionary literature to model gradual adjustment and learning in games (see, for example, Blume (1993), Durlauf (1997), or Young (1998)). It is in the spirit of the well-known formulation of *logistic quantal response equilibrium* proposed by McKelvey and Palfrey (1995), which has been provided with a natural bounded-rationality interpretation by Chen et al. (1997). The parameter β modulates the noise impinging on agents' adjustment. If $\beta = 0$, noise is the overwhelming force and all actions are chosen with the same probability, irrespectively of local conditions and payoffs. In contrast, in the polar case where β is very large, only if a particular action is a genuine best response is it chosen with a sizable probability.

8.3 Analysis

For finite β, the model is substantially more complex than the degenerate version studied in MEV. Consequently, we are unable to obtain an exact characterization of its stable equilibria and thus have to base our analysis on some simplifying assumptions. Naturally, this implies that the solution we arrive at can no longer be regarded as a fully accurate description of the long-run behaviour of the model. The entailed approximation, however, turns out to be quite effective since, as we shall explain, it matches very closely the results obtained from numerical simulations for large populations.

The adjustment rule given by (8.1) not only has the precedents in economics summarized in Sect. 8.2 but is also formally identical to the "spin dynamics" postulated by the so-called Potts model in statistical physics, itself a generalization of the canonical Ising model. In this context, $1/\beta$ plays the role of the temperature at which the particle interaction takes place and the q different actions are the possible spins. (See, for example, Vega-Redondo (2007) for a detailed explanation of these models and their relationship to the economic and evolutionary literature.) The Potts model has been studied in detail by physicists and exact solutions for it exist for low-dimensional lattices as well as trees (cf. Baxter (1982)). Recently, the analysis has been extended to random networks by Dorogovtsev et al. (2004) and Ehrhardt and Marsili (2005). We crucially rely on the latter in our present analysis.

One of the simplifying assumptions we make is that the network prevailing at any given point in time is a random network suitably characterized by a degree distribution

$$\boldsymbol{p} \equiv \{p(k)\}_{k=0}^{\infty}$$

that specifies the fraction of nodes *p(k)* that display each possible degree k. The defining property of a *random network* is the absence of statistical correlations. Thus, in particular, it is presumed that the degree of a node is stochastically independent of any of its neighbours. Such a property does not strictly hold in our present generalized context – only approximately so.[2] This is why the random-network postulate must be viewed, in this case, as a convenient, but not fully accurate, description of the system at any point in time.

Another simplifying assumption we shall make concerns the relative speed of action and link adjustment. For technical tractability (and in particular, to rely on the solutions of the Potts model available in the literature), it is convenient to posit that the network adjusts much slower than agents' actions. Formally, this is captured by making $v \to \infty$. It amounts to assuming that actions adjust at a much brisker pace than links, so that the current underlying network may be taken as fixed while the action distribution reaches a stable configuration.

Under these assumptions, the analysis of the model can be decomposed into the following steps. First, we need to derive the law of motion for the degree distribution **p**, which is governed by the subprocesses of link creation and link destruction. As explained, the latter simply has every existing link disappear at a constant rate λ. Link creation, on the other hand, depends on the probability that any two agents who meet and have the potential of creating a new link happen to display the same action. For any given agent/node i, this ex ante probability must generally depend on its degree $z_i = k$, so we denote it by $\pi(k)$. In essence, this probability results from the combination of the following three constituent probabilities:

1. The (unconditional) probability ζ that, when node i selects another node at random, the latter belongs to the (unique)[3] giant component of the network.
2. The conditional probability $\gamma(k)$ that node i of degree k belongs to the giant component.

[2]By way of illustration, one of the features of the process that introduces internode correlations can be explained as follows. First note that the postulated action dynamics leads high-degree nodes to exhibit, on average, stronger social "conformity" than lower-degree nodes. That is, they have a higher probability of choosing the action that is in the majority in the population. This in turn implies that links between high degree nodes will be formed with higher probability (that is, at a higher rate) than between lower-degree nodes. In the end, therefore, positive degree correlations will tend to arise, high-degree nodes being more likely to be connected to other high-degree nodes than what is prescribed by the *unconditional* average.

[3]As is well known in the theory of random networks, if a giant component exists in this context, it is unique.

3. The conditional probability $\mu(k)$ that, if node i of degree k does belong to the giant component, its action coincides with that of a randomly selected node in that component.

The first probabilities, ζ and $y(k)$ for each k, only depend on the underlying degree distribution $\mathbf{p}$. The probabilities $\mu(k)$, on the other hand, depend both on $\mathbf{p}$ and the value of β in (8.1) that modulates the pressure towards local conformity induced by the action dynamics. To compute the probabilities in (1)–(2), one can directly use the standard techniques of the modern theory of random networks, as explained, for example, in Vega-Redondo (2007). And concerning the probability in (3), we may rely on the aforementioned solution of the Potts model in random networks that has been developed by Ehrhardt and Marsili (2005).

Thus let ζ, $\gamma(k)$ and $\mu(k)$ for each k be the probabilities prevailing at some point in time when the underlying network is modelled as a random network with degree distribution $\mathbf{p}$. Then, the probability $\pi(k)$ that a randomly chosen node of degree k meets a node that displays its own action is simply given by:

$$\pi(k) = (1 - \zeta\gamma(k))\frac{1}{q} + \zeta\gamma(k)\mu(k). \tag{8.2}$$

The second term in the right-hand side of (8.2) corresponds to the event that some arbitrary node i of degree k happens to be part of the giant component and it meets another node j in that same component. In that case, the link between i and j is formed with probability $\mu(k)$, that is, the probability that both display the same action.[4] The first term, on the other hand, contemplates what happens when either node i or/and the node j it meets are *not* in the giant component. Then, with probability essentially one in a large random network, both nodes are in different components. This implies that they will only display the same action (and thus form a link) "by chance", that is, with probability $1/q$ since there are q possible actions.

Given the probabilities $\pi(k)$ specified in (8.2), the evolution of the degree distribution in time can be modelled through the following differential equation:

$$\dot{p}(k) = (k+1)\lambda p(k+1) + 2\eta p(k-1)\pi(k-1) - k\lambda p(k) - 2\eta p(k)\pi(k), \tag{8.3}$$

where we dispense with the time index for notational simplicity. The first two terms in the right-hand side of (8.3) reflect the inflow into the frequency of nodes of degree k. This inflow consists of those nodes that had degree $k+1$ and lost one of its links (which happens at a rate λ per link), combined with the rate at which nodes with degree $k-1$ form a new link. (In the latter respect, note that the factor of 2 accounts for the fact that a link is created if either the node in question receives the initiating opportunity or some other node does.) On the other hand, the two last terms embody the opposite flow that decreases the frequency of nodes of degree k when these nodes either loose or create a link.

[4]Of course, this presumes that the link between i and j is not already in place, which is an event that can be essentially ignored in large populations.

We are interested in characterizing the pair

$$\mathbf{p}^* = \{p^*(k)\}_{k=0}^{\infty}, \pi^* = \{\pi^*(k)\}_{k=0}^{\infty}$$

that defines a stationary point of the dynamical system. Such a stationarity embodies a twin requirement. First, given $\mathbf{p}^*$, the corresponding π^* induced by (8.2) must suitably characterize the long-run coordination probabilities induced by the fast dynamics governed by (8.1). Second, given π^*, the degree distribution $\mathbf{p}^*$ must define a stationary point of (8.3). This latter requirement simply amounts to stating that, for all $k = 1, 2, \ldots$,

$$(k+1)\lambda p^*(k+1) + 2\eta p^*(k-1)\pi^*(k-1) = k\lambda p^*(k) + 2\eta p^*(k)\pi^*(k).$$

This defines a system of difference equations that can be solved recursively as follows:

$$p^*(1) = 2\eta p^*(0)\pi^*(0), \tag{8.4}$$

along with

$$p^*(k+1) = \frac{2\eta p^*(k)\pi^*(k) + k\lambda p^*(k) - 2\eta p^*(k-1)\pi^*(k-1)}{\lambda(k+1)} \quad (k = 1, 2, \ldots), \tag{8.5}$$

once we impose the normalization

$$\Sigma_{k=0}^{\infty} p^*(k) = 1.$$

To understand the essential gist of the argument, it is useful to make the assumption that the underlying degree distribution $\mathbf{p}$ is Poisson. (This assumption is not exactly true, but happens to be quite a good approximation for most parameter values in the interesting range.) Then, we can define the average probability $\pi(c, \beta)$ that two randomly selected nodes happen to be well coordinated (and thus will establish a link), as a function of β (the noise parameter) and the average degree c (the only parameter characterizing a Poisson degree distribution). The desired stationarity of the situation then requires that average link destruction be equal to average link creation, that is,

$$\lambda c = \pi(c, \beta). \tag{8.6}$$

The situation is described in Fig. 8.1 for different values of λ, while we normalize the rate of link creation η to unity. It shows that when $\lambda > \lambda_2$ there is a single solution, representing a sparse network. At λ_2 other two solutions arise, one of which is unstable. At a further point $\lambda_1 < \lambda_2$ the sparse-network solution merges

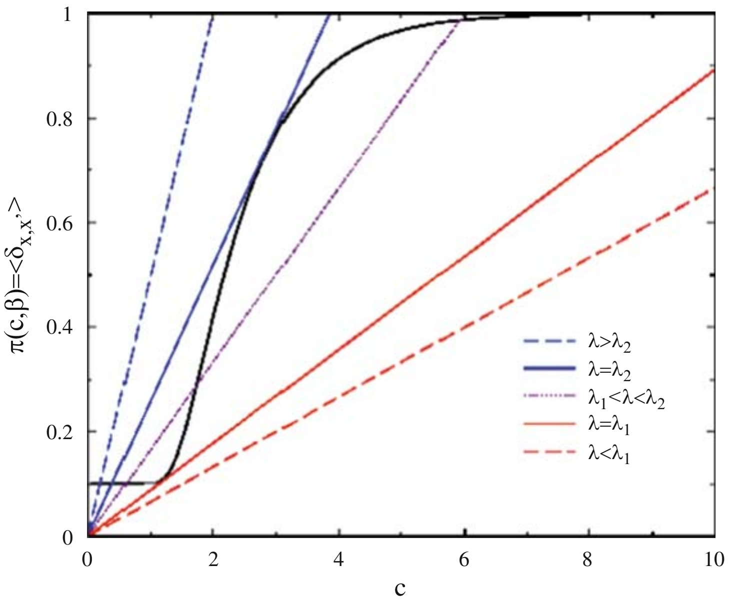

Fig. 8.1 Graphical illustration of the solutions for the stationarity condition (8.6) for $q = 10$ and $\beta = 4$, under the assumption that the underlying random network is Poisson. The solutions are given by intersections of the curve representing the function $\pi(c; \beta)$ with rays λc for different values of λ. For $\lambda = \lambda_1$ and $\lambda = \lambda_2$ the rays are tangent to the curve, thus marking the values that bound the region $\lambda \in [\lambda_1, \lambda_2]$ where multiple intersections (that is, solutions) exist. Outside this region, (8.6) displays a unique solution

with the unstable one and both disappear for $\lambda < \lambda_1$, leaving only a solution with a stable and dense network. This reproduces the phenomenology summarized in (1)–(3) in Sect. 8.1 and, as explained, coincides fully with that obtained in EMV for the noiseless degenerate model.[5]

To conclude, we focus our discussion on the role played by the new parameter β that marks the only difference with the EMV model. Interestingly, we find that changes in β induce the same qualitative pattern of long-run behaviour as observed before for changes in η. The conclusions – both theoretical and simulations – are depicted in Fig. 8.2 for two different values of η. (In this case, we find it convenient to scale time by normalizing the rate λ to unity.)

Figure 8.2 shows that, as one implements gradual changes in the noise impinging on action adjustment (which can be suitably parameterized by $1/\beta$), the long-run

[5]Mathematically, the behavior displayed by the model reflects the onset of a bifurcation towards instability and equilibrium multiplicity as the parameter λ enters the region $[\lambda_1, \lambda_2]$. Such a bifurcation is analogous to that found by Brock and Hommes (1997) as the intensity of choice (here captured by either a change in the volatility rate λ or the choice-sensitivity parameter β) varies in a suitable range. See also Hommes (2006) for an extensive discussion of the issue.

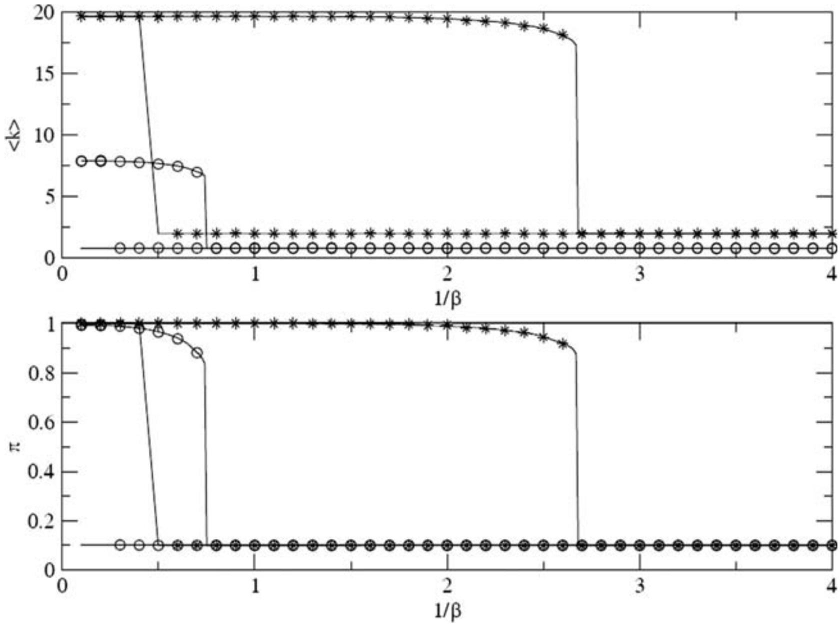

Fig. 8.2 The *upper panel* plots the average connectivity $\langle \mathrm{k} \rangle$ predicted by the model against the noise level $1/\beta$, for $q=10$ and two different values of η, that is, $\eta=4$ (*lower curves*), $\eta=10$ (*higher curves*). The *solid lines* trace the theoretical prediction while the points represent simulation results for $n = 1{,}000$.The *lower panel* displays analogous results for the average probability π that two randomly chosen nodes display the same action

behaviour of the system displays the same three features that are obtained for changes in η That is, both connectivity as well as social conformity exhibit sharp and resilient transitions across multiple equilibria as the noise level changes "slightly" around certain thresholds. It is worth stressing that the theoretical predictions are well supported by numerical simulations, even though, as we have explained, the analytically solved model can only be conceived as an approximate description of the system dynamics.

8.4 Conclusion

We conclude, therefore, that the introduction of noise into the model studied by MEV maintains the essential phenomenology that was encountered there for the degenerate noiseless version of the model, that is, sharp transitions, hysteresis, and equilibrium multiplicity are robust features of the long-run dynamics of the process when the link-destruction or link-creation rates change in a relevant range. In fact, we have found that analogous behaviour arises as well concerning changes in the parameter controlling for the action-revision noise on which we have focused our analysis here. This suggests that such a phenomenology may well represent a solid (and, in a sense, universal) pattern to be expected in network formation processes reflecting the forces of homophily and conformity.

Building upon those insights, there are two different avenues we want to explore in future research. First, we would like to understand how the model fares when the mechanism of link creation is endowed with a natural local dimension – for example, if new linking opportunities are assumed to be found through the "intermediation" of current neighbours, as in Marsili et al. (2004). Second, we would like to enrich the present abstract modelling approach with features (say, payoffs and purposeful decision making) that would allow one to study important and concrete economic problems, for example, the role of R&D collaboration on firm innovation in oligopolistic industries.

Acknowledgments We thank an anonymous referee for very useful comments. Vega-Redondo also acknowledges financial support from the Spanish Ministry of Education under grant SEJ2007–62656.

References

Baxter RJ (1982) Exactly solved models in statistical mechanics. Academic Press, London

Blume L (1993) The statistical mechanics of strategic interaction. Games Econ Behav 4:387–424

Brock WA, Hommes CH (1997) A rational route to randomness. Econometrica 65:1059–1095

Calvó-Armengol A, Jackson MO (2004) The effects of social networks on employment and inequality. Am Econ Rev 94:426–454

Castells M (1996) The information age: economy, society, and culture, volume I: The rise of the network society. Blackwell, Massachusetts

Chen H-C, Friedman JW, Thisse J-F (1997) Boundedly rational Nash Equilibrium: a probabilistic choice approach. Games Econ Behav 18:32–54

Dorogovtsev SN, Goltsev AV, Mendes JFF (2004) Potts model on complex networks. Eur Phys J B 38:177–182

Durlauf S (1997) Statistical mechanics approaches to socioeconomic behavior. In: Arthur WB, Durlauf SN, Lane DA (eds) The economy as an evolving complex system II. Addison-Wesley, Reading, MA, pp 81–104

Ehrhardt GCMA, Marsili M (2005) Potts model on random trees. J Stat Mech P02006

Ehrhardt GCMA, Marsili M, Vega-Redondo F (2006) Networks emerging in a volatile world. Preprint Abdus Salam Center for Theoretical Physics and Universidad de Alicante

Granovetter M (1974) Getting a job: a study on contacts and careers. Chicago University Press, Chicago

Hagedoorn J (2002) Inter-firm R&D partnerships: an overview of major trends and patterns since 1960. Res Pol 31:477–492

Hommes CH (2006) Heterogeneous agent models in economics and finance. In: Tesfatsion L, Judd KJ (eds) Handbook of computational economics, vol 2: Agent-based computational economics. Elsevier, Amsterdam, pp 1109–1186

Kirman A, Herreiner DK, Weisbuch G (2000) Market organization and trading relationships. Econ J 110:411–436

Kranton R, Minehart D (2001) A theory of buyer-seller networks. Am Econ Rev 91:485–508

Marsili M, Vega-Redondo F, Slanina F (2004) The rise and fall of a networked society: a formal model. Proc Natl Acad Sci USA 101:1439–1442

McKelvey RD, Palfrey TR (1995) Quantal response equilibria for normal form games. Games Econ Behav 10:6–38

Montgomery J (1991) Social networks and labor market outcomes: toward an economic analysis. Am Econ Rev 81:1408–1418

Powell WW, White DR, Koput KW et al. (2005) Network dynamics and field evolution: the growth of inter-organizational collaboration in the life sciences. Am J Sociol 110:1132–1205

Vega-Redondo F (2007) Complex social networks, econometric society monograph series. Cambridge University Press

Young P (1998) Individual strategy and social structure: an evolutionary theory of institutions. Princeton University Press, Princeton, NJ