DISCUSSION PAPER SERIES

No. 8807

NESTEDNESS IN NETWORKS: A THEORETICAL MODEL AND SOME APPLICATIONS

Michael König, Claudio J. Tessone and Yves Zenou

PUBLIC POLICY

Centre for Economic Policy Research

www.cepr.org

Available online at: www.cepr.org/pubs/dps/DP8807.asp

ISSN 0265-8003

NESTEDNESS IN NETWORKS: A THEORETICAL MODEL AND SOME APPLICATIONS

Michael König, Stanford University
Claudio J. Tessone, ETH Zurich
Yves Zenou, Stockholm University

Discussion Paper No. 8807
January 2012

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 8807

January 2012

ABSTRACT

Nestedness in Networks: A Theoretical Model and Some Applications*

We develop a dynamic network formation model that can explain the observed nestedness in real-world networks. Links are formed on the basis of agents' centrality and have an exponentially distributed life time. We use stochastic stability to identify the networks to which the network formation process converges and find that they are nested split graphs. We completely determine the topological properties of the stochastically stable networks and show that they match features exhibited by real-world networks. Using four different network datasets, we empirically test our model and show that it fits well the observed networks.

JEL Classification: A14, C63 and D85
Keywords: Bonacich centrality, nested split graphs, nestedness and network formation

Michael König
SIEPR and Department of Economics,
Stanford University,
579 Serra Mall, CA 94305-6072, USA
Email: mdkoenig@stanford.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=171073

Claudio J. Tessone
Chair of Systems Design,
Department of Management,
Technology and Economics, ETH Zurich, Kreuzplatz 5,
CH-8032 Zurich,
Switzerland
Email: tessonec@ethz.ch

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=171075

Yves Zenou
Department of Economics
Stockholm University
10691 Stockholm
SWEDEN
Email: yves.zenou@ne.su.se

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=126027

* We would like to thank Phillip Bonacich, Yann Bramoullé, Ulrik Brandes, Guido Caldarelli, Sanjeev Goyal, Patrick Groeber, Matt Jackson, Matteo Marsili, Fernando Vega-Redondo, Douglas White, Eric Gilson as well as the participants of the 2009 DIME conference in Paris, the 2009 Trento Summer School on "Innovation and Networks", the 2009 HSC Videoconference at the University of California San Diego, the 2009 workshop on the Economics of Social Networks at Laval University, the Microeconomic Theory seminar at University of California Berkeley in 2009, the Theory Workshop at the Kellog School of Management (MEDS) at Northwestern University in 2010 and the 2011 Society for Advanced Economic Theory (SAET) conference in Faro for their helpful comments. We would like to thank the Austrian National Bank for providing access to the data. Michael D. König acknowledges financial support from Swiss National Science Foundation through research grant PBEZP1–131169.

Submitted 12 January 2012

Nestedness in Networks: A Theoretical Model and Some Applications ☆

Michael D. König[a], Claudio J. Tessone[b], Yves Zenou[c]

[a] *SIEPR and Department of Economics, Stanford University, 579 Serra Mall, CA 94305-6072, United States.*
[b] *Chair of Systems Design, Department of Management, Technology and Economics, ETH Zurich, Kreuzplatz 5, CH-8032 Zurich, Switzerland*
[c] *Stockholm University and Research Institute of Industrial Economics (IFN). Sweden.*

Abstract

We develop a dynamic network formation model that can explain the observed nestedness in real-world networks. Links are formed on the basis of agents' centrality and have an exponentially distributed life time. We use stochastic stability to identify the networks to which the network formation process converges and find that they are nested split graphs. We completely determine the topological properties of the stochastically stable networks and show that they match features exhibited by real-world networks. Using four different network datasets, we empirically test our model and show that it fits well the observed networks.

Key words: Nestedness, Bonacich centrality, network formation, nested split graphs
JEL: A14, C63, D85

1. Introduction

Nestedness is an important aspect of real-world networks.[1] For example, the organization of the New York garment industry [Uzzi, 1996] and of the Fedwire bank network [May et al., 2008; Soramaki et al., 2007] are nested in the sense that their organization is strongly hierarchical. If we consider the latter network (i.e. the network of banks), then as reported in May et al. [2008], the topology of interbank payment flows within the US

☆We would like to thank Phillip Bonacich, Yann Bramoullé, Ulrik Brandes, Guido Caldarelli, Sanjeev Goyal, Patrick Groeber, Matt Jackson, Matteo Marsili, Fernando Vega-Redondo, Douglas White, Eric Gilson as well as the participants of the 2009 DIME conference in Paris, the 2009 Trento Summer School on "Innovation and Networks", the 2009 HSC Videoconference at the University of California San Diego, the 2009 workshop on the Economics of Social Networks at Laval University, the Microeconomic Theory seminar at University of California Berkeley in 2009, the Theory Workshop at the Kellog School of Management (MEDS) at Northwestern University in 2010 and the 2011 Society for Advanced Economic Theory (SAET) conference in Faro for their helpful comments. We would like to thank the Austrian National Bank for providing access to the data. Michael D. König acknowledges financial support from Swiss National Science Foundation through research grant PBEZP1–131169.

Email addresses: mdkoenig@stanford.edu (Michael D. König), tessonec@ethz.ch (Claudio J. Tessone), yves.zenou@ne.su.se (Yves Zenou)

[1] We speak of nestedness in terms of a nested neighborhood structure of the nodes in a network: A network exhibits *nestedness* if the neighborhood of a node is contained in the neighborhoods of the nodes with higher degrees.

Fedwire service is clearly nested.[2] The sample from this network amounted to around $700,000$ transfer funds, with just over $5,000$ banks involved on an average day. The authors [Soramaki et al., 2007] find that this network is characterized by a relatively small number of strong flows (many transfers) between nodes, with the vast majority of linkages being weak to non-existing (few to no flows). Furthermore, the topology of this Fedwire network is highly *dissortative* [Newman, 2002], since large banks are disproportionately connected to small banks, and vice versa; the average bank was connected to 15 others. In other words, most banks have only a few connections while a small number of "hubs" have thousands. Similarly, Åkerman and Larsson [2010], who study the evolution of the global arms trade network using a unique dataset on all international transfers of major conventional weapons over the period 1950-2007, find that these networks are nested and dissortative in the sense that big countries mainly trade arms with small countries but small countries do not trade with each other. Using aggregate bilateral imports from 1950 to 2000, De Benedictis and Tajoli [2011] analyze the structure of the world trade network over time, detecting and interpreting patterns of trade ties among countries. They also find that world trade tends to be concentrated among a sub-group of countries and a small percentage of the total number of flows accounts for a disproportionately large share of world trade. The larger countries account for a generally larger share of world trade and have more partners. Figure 3 in their paper shows a clear core-periphery structure, indicating nestedness of their networks. Interestingly, in all these networks, dissortativity arises naturally since "big" agents tend to interact with "small" agents and vice versa. For example, banks seek relationships with each other that are mutually beneficial. As a result, small banks interact with large banks for security, lower liquidity risk and lower servicing costs, and large banks may interact with small banks in part because they can extract a higher premium for services and can accommodate more risk.

Surprisingly, nestedness has not been studied from a theoretical point of view, even though other salient features of networks such as "small world" properties with high clustering and short average path lengths [Watts and Strogatz, 1998] as well as "scale-free" or power-law degree distribution [Barabasi and Albert, 1999] have received a lot of attention.[3]

The first aim of this paper is to propose a dynamic network formation model that exhibits not only the standard features of real-world networks (small worlds, high clustering, short path lengths and a power-law degree distributions) but also *nestedness* and *dissortativity*. The second aim is to provide a microfoundation for the network formation process where linking decisions are based on the utility maximization of each agent rather than on a random process, which is often assumed in most dynamic models of network formation. The last aim of this paper is to test empirically our model with different network data sources (interbank loans, trade in conventional goods and arms trade between countries) and to show that our model fits well the features of these real-world networks, especially their nestedness.

To be more precise, we develop a dynamic model where, at each period of time, agents play a two-stage game: in the first stage, as in Ballester et al. [2006], agents play their equilibrium contributions proportional to their *Bonacich centrality* [Bonacich, 1987], while, in the second stage, a randomly chosen agent can update her linking strategy by creating a new link as a best response to the current network. Links do not last forever but have an

[2]See Figures 1 and 2 in Soramaki et al. [2007].

[3]See Jackson and Rogers [2007] who propose a model that has all these features but not nestedness.

exponentially distributed life time. The most valuable links (i.e. the ones with the highest Bonacich centrality) decay at a lower rate than those that are less valuable. As a result, the formation of social networks can be regarded as a tension between the search for new linking opportunities and volatility that leads to the decay of existing links.

We introduce noise into the decision process to form links [cf. Blume, 2003; Feri, 2007; Foster and Young, 1990; Hofbauer and Sandholm, 2007; Kandori et al., 1993; Sandholm, 2010], and analyze the limit of the invariant distribution, the stochastically stable networks, as the noise vanishes. We first show that in this limit, starting from arbitrary initial conditions, at each period of time the network generated by this dynamic process is a nested split graph. These graphs, which are relatively well-known in the applied mathematics literature [Aouchiche et al., 2008; Mahadev and Peled, 1995], have a very nice and simple structure that make them very tractable to work with. To the best of our knowledge, this is the first time that a complex dynamic network formation model can be characterized by such a simple structure in terms of networks it generates. By doing so, we are able to bridge the economics literature and the applied mathematics/physics literatures in a simple way. Because of their simple features, we then show that degree, closeness, eigenvector and Bonacich centrality induce the same ordering of nodes in a nested split graph (this is also true for betweenness centrality if the ordering is not strict). This implies, in particular, that if we had a game where agents formed links according to other measures of centrality (such as degree, closeness, or betweenness) than the Bonacich centrality, then all our results would be unchanged. We then show that the stochastically stable network is a nested split graph. Instead of relying on a mean-field approximation of the degree distribution and related measures as most dynamic network formation models do, because of the nature of nested split graphs, we are able to derive explicit solutions for all network statistics of the stochastically stable networks (by computing the adjacency matrix).[4] We also find that, by altering the rate at which linking opportunities arrive and links decay, a sharp transition takes place in the network density. This transition entails a crossover from highly centralized networks when the linking opportunities are rare and the link decay is high to highly decentralized networks when many linking opportunities arrive and only few links are removed. From the efficiency perspective such sharp transition can also be observed in aggregate payoffs in the stochastically stable networks.

The intuition of these results is as follows. Agents want to link to other agents who are more central since this leads to higher actions (as actions are proportional to centrality) and higher actions raise payoffs more. Similarly, links to agents with lower centrality last shorter. Notice moreover that, once someone loses a link with an agent, she becomes less central and this makes it more likely that the next link she has will also disappear. Thus link gains and losses are self reinforcing. This intuition suggests that if α, the probability of adding links is large then the process should approximate complete network while if it is small then the process should approximate the star network. The key insight of our model is that *for intermediate values of α the stochastically stable network is a nested split graph*.

We then proceed by showing that our model reproduces the main empirical observations of social networks. Indeed, we show that the stochastically stable networks emerging in our link formation process are characterized by *short path length* with *high clustering*, *exponential degree distributions* with *power-law tails*, *negative degree-clustering correlation*

[4] In a nested split graph the degree distribution uniquely defines the adjacency matrix (up to a permutation of the node labels).

and *nestedness*. These networks also show a clear core-periphery structure. Moreover, we show that stochastically stable networks are *dissortative*.

Using four different data sources we then test empirically our model. The first network we analyze is the network of Austrian banks in the year 2008. Links in the network represent exposures between Austrian-domiciled banks on a non-consolidated basis [cf. Boss et al., 2004]. The second network is the global banking network in the year 2011 obtained from the Bank of International Settlements (BIS) locational statistics on exchange-rate adjusted changes in cross-border bank claims [cf. Minoiu and Reyes, 2011]. The third one is a trade network between countries in the year 2000. The trade network is defined as the network of import-export relationships between countries in a given year in millions of current-year U.S. dollars. Finally, we consider the network of arms trade between countries [cf. Åkerman and Larsson, 2010]. We use data obtained from the SIPRI Arms Transfers Database holding information on all international transfers between countries of seven categories of major conventional weapons accumulated from 1950 to 2010. A link in the network represents a recipient or supply relationship between two countries during this period. Even though these networks are very different, they all exhibit strong nestedness and dissortativity and we find a reasonable goodness of fit of our model with these networks (even though our model is only parsimoniously parameterized).

It is worth noting that the model we study is rather general and the fact that our network formation process generates nested-split graphs is not an artifact of a very specific linking protocol. First, our results are independent of initial conditions. That is, the network formation process will always converge to a nested split graph starting from any possible initial network eventually. Second, we would obtain exactly the same results using more general utility functions, e.g. any increasing function of the Bonacich centrality of the agent or using "information centrality" introduced in Stephenson and Zelen [1989]. Third, when creating a link, an agent can choose anyone in the network not only neighbors of neighbors as it is often assumed. In other words, we do not restrict the set of choices of her possible contacts. Fourth, if we relax the assumptions that agents are myopic and consider instead farsighted agents, we can still show that a stochasically stable network with nested-split feature exists. Finally, if it is assumed that link formation is costly, we are able to show that if this cost is not too large, networks will still be nested split graphs. We discuss in more detail all these robustness issues in Section 9.

Our paper is organized as follows. Section 2 discussed the relation of our model to the literature. In Section 3, we introduce the model and discuss the basic properties of the network formation process. In particular, Section 3.1 discusses the first stage of the game. In Section 3.2, we introduce the second stage of the game, where the network formation is explained. Next, Section 4 shows that stochastically stable networks exist, can be computed analytically and are nested-split graphs. After deriving the stochastically stable networks in Section 5, we analyze their properties in terms of topology and centralization. In Section 6, we study efficiency from the point of view of maximizing total efforts and aggregate payoff in the stochastically stable networks. We investigate the efficiency of different stochastically stable networks as a function of the volatility of the environment. In Section 7, we show that the particular structure of nested-split graphs can be observed in real-world networks and run some numerical simulations to show the resilience of nested-split graphs to changes in parameters. Using four different network datasets, we empirically test our model in Section 8. Section 9 discusses our results and their robustness. Finally, Section 10 concludes. Appendix A gives all the necessary definitions and characterizations of networks used throughout the paper. In Appendix B we provide some general results for nested split graphs in terms of their topology properties and centralization measures. We extend our analysis in Appendix C by including linking costs. All proofs can be found in Appendix D.

2. Relation to the literature

The literature on network formation is basically divided in two strands that are not communicating very much with each other. In the random network approach (mainly developed by mathematicians and physicists),[5] which is mainly dynamic, the reason why a link is formed is pure chance. Indeed, this literature builds networks either through a purely stochastic process where links appear at random according to some distribution, or else through some algorithm for building links. In this approach, researchers study how emerging networks match real-world networks [see e.g. Newman, 2010; Vega-Redondo, 2007]. While sharing some common features with this literature, our model is quite different since agents do not create links randomly but in a strategic way, i.e. they maximize their utility function.

In the other approach (developed by economists), the reason for the formation of a link is strategic interactions. Individuals carefully decide with whom to interact and this decision entails some consent by both parts in a given relationship. There are some dynamic network formation models with strategic interactions. Bala and Goyal [2000], Watts [2001], Jackson and Watts [2002a], Dutta et al. [2005] are prominent papers of this literature. Bala and Goyal [2000] propose a model similar to the connections model [Jackson and Wolinsky, 1996], but with directional flow of communication or information. They focus on the dynamic formation of networks in this directed communications model. The network formation game is played repeatedly, with individuals deciding on link formation in each period. In this setting, for low enough costs to forming links, the process leads naturally to a limiting network which has the efficient structure of a wheel. Also in the context of the connections model, Watts [2001] considers a framework where pairs of agents meet over time, and decide whether or not to form or sever links with each other. Agents are myopic and so base their decision on how the decision on the given link affects their payoffs, given the current network in place. An interesting result applies to a cost range where a star network is both pairwise stable and efficient. Jackson and Watts [2002b] model network formation as an intertemporal process with myopic individuals breaking and forming links as the network evolves dynamically. Dutta et al. [2005] relax the assumption of myopic agents and assume that agents behave in a farsighted manner by taking into account the intertemporal repercussions of their own decisions. Our model is different than the ones developed in these papers in the sense that we consider both dynamic models of network formation and optimal actions from agents. This allows us to give a microfoundation of the network formation process as equilibrium actions transform into equilibrium utility functions. Another crucial difference is that we are able to match most features of real-world networks while these models do not.[6] As Jackson [2007, 2008] pointed out, the

[5] See e.g. Albert and Barabási [2002].

[6] Mele [2010] provides an interesting dynamic network formation model where individuals decide with whom to form links by maximizing a utility function. Even though the utility function has nice properties (it is a potential function), contrary to our model, Mele does not characterize analytically the network statistics we do here and resorts to Bayesian estimation strategy to estimate his model.

random approach gives us a great deal of insight into *how* networks form (i.e. matches the characteristics of real-life networks) while the deterministic approach performs better on *why* networks form.

There is also another strand of the literature (called "games on networks") that takes the network as given and studies how the network structure impacts on outcomes and individual decisions. A prominent paper of this literature is Ballester et al. [2006].[7] They mainly show that if agents' payoffs are linear-quadratic, then the unique interior Nash equilibrium of an $n-$player game in which agents are embedded in a network is such that each individual effort is proportional to her Bonacich centrality measure. The latter is a well-known centrality measure introduced by Bonacich [1987].[8] The Bonacich centrality of a particular node counts the total number of paths that start from this node in the graph, weighted by a decay factor based on path length. In the model of Ballester et al. [2006], it is mainly the centrality of an agent in a network that explains her outcome.[9] In this paper, we introduce *strategic interactions* in a non-random dynamic network formation game where agents also choose how much effort they put in their activities.

There are some papers that, as in our framework, combine both network formation and endogenous actions. These papers include Bramoullé et al. [2004], Cabrales et al. [2010], Calvo-Armengol and Zenou [2004], Galeotti and Goyal [2010], Goyal and Vega-Redondo [2005], Goyal and Joshi [2003], Jackson and Watts [2002a]. Most of these models are, however, static and the network formation process is different. Goyal and Joshi [2003] is the closest to our model. They consider a standard Cournot model of competition where prior to market stage, firms can form costly links with each other since a link lowers costs of production for the two firms. In this framework, a network defines a cost profile for the competing firms. They show that every pairwise equilibrium connected network is an inter-linked star network. Their model is static but has link formation and choice of action (quantity). Our model is dynamic and has link formation and a choice of action. The details of the payoffs and the methods of analysis, however, differ. In particular, the introduction of stochastic elements in our analysis helps in solving our model, which allows us to obtain more general results, like for example calculating the exact degree distribution of the stochastically stable networks.

Finally, our paper is also related to Jackson and Rogers [2007], who also motivate their modelling approach by means of statistics of empirical networks. In their model, new nodes are born over time and can attach to existing nodes either by choosing one uniformly at random or through a search process that makes the likelihood of meeting a given node

[7]Bramoullé and Kranton [2007], Bramoullé et al. [2011] and Galeotti et al. [2010] are also important papers in this literature. The first paper focuses on strategic substitutabilities while the second one provides a general framework solving any game on networks with perfect information and linear best-reply functions. The last paper investigates the case when agents do not have perfect information about the network. Because of its tractability, in the present paper, we use the model of Ballester et al. [2006] who analyze a network game of local complementarities under perfect information.

[8]Centrality is a fundamental measure of the importance of actors in social networks, dating back to early works such as Bavelas [1948]. See Wasserman and Faust [1994] for an introduction and survey.

[9]In the empirical literature, it has been shown that centrality is important in explaining exchange networks [Cook et al., 1983], peer effects [Calvó-Armengol et al., 2009; Haynie, 2001], creativity of workers [Perry-Smith and Shalley, 2003], workers' performance [Mehra et al., 2001], power in organizations [Brass, 1984], the flow of information [Borgatti, 2005; Stephenson and Zelen, 1989], the formation and performance of R&D collaborating firms and inter-organizational networks [Boje and Whetten, 1981; Powell et al., 1996; Uzzi, 1997] as well as the success of open-source projects [Grewal et al., 2006].

proportional to the number of links the node already has. In their model, m is the average degree while r represents the ratio of how many links are formed uniformly at random compared to how many are formed proportionally to the number of links existing nodes already have. As r approaches 0, the distribution converges to be scale-free, while as m tends to infinity the distribution converges to a negative exponential distribution, which corresponds to the degree distribution of a purely uniform and independent link formation process on a network that grows over time. Our model is quite different since, contrary to Jackson and Rogers [2007], agents choose actions and form links by maximizing their utility. Also we look at the steady-state distribution while they analyze growing networks. As a result, the predictions of the two models are quite different. In particular, we can explain how and why nestedness is an important feature of real-world networks.[10]

To summarize, our main contribution to the literature is that we are able to explain the emergence of nestedness in networks by analyzing a dynamic network formation model with endogenous actions. We are also able to analytically characterize the stochastically stable networks, which can be shown to be nested-split graphs, and to provide a microfoundation for the link-formation process. Even if nested-split graphs have a much more regular structure than the complex networks we observe in many real-world applications, they are easy to study, they are the result of endogenous rational actions and they have most of the properties of real-world networks. Finally, we test empirically our model with four different datasets and show that our model fits well these observed networks.

3. The Model

In this section, we introduce the network formation process, which can be viewed as a two-stage game on two separate time scales. On the fast time scale, all agents simultaneously choose their effort level in a fixed network structure. It is a game following Ballester et al. [2006] with local complementarities where players have linear-quadratic payoff functions. On the slow time scale, agents receive linking opportunities at a given rate and decide with whom they want to form a link while the links they have created decay after having reached their finite life time. This introduces two different time scales, one in which agents are choosing their efforts in a simultaneous move game and the second in which an agent forms a link and anticipates the equilibrium outcome in the following simultaneous move game.

3.1. Nash Equilibrium and Bonacich Centrality

Consider a static network G in which the nodes represent a set $\mathcal{N} = \{1, 2, \ldots, n\}$ of agents/players. Following Ballester et al. [2006], each agent $i \in \mathcal{N}$ in the network G selects an effort level $x_i \geq 0$, $\mathbf{x} \in \mathbb{R}_+^n$, and receives a payoff $\pi_i : \mathbb{R}_+^n \times \Omega \times \mathbb{R}_+ \to \mathbb{R}$ of the following form

$$\pi_i(\mathbf{x}, G, \lambda) \equiv x_i - \frac{1}{2}x_i^2 + \lambda \sum_{j=1}^{n} a_{ij} x_i x_j, \tag{1}$$

[10] Also, the model by Jackson and Rogers [2007] produces assortative networks, while ours generates dissortative networks. Hence, the real-world networks to which these two models apply are different.

where $\lambda \geq 0$ and $a_{ij} \in \{0,1\}$, $i,j = 1,\ldots,n$ are the elements of the symmetric $n \times n$ *adjacency matrix* $\mathbf{A}$ of G. This utility function is additively separable in the idiosyncratic effort component $(x_i - 1/2x_i^2)$ and the peer effect contribution $(\lambda \sum_{j=1}^n a_{ij}x_ix_j)$. Payoffs display strategic complementarities in effort levels, i.e., $\partial^2\pi_i(\mathbf{x}, G, \lambda)/\partial x_i \partial x_j = \lambda a_{ij} \geq 0$.

The general payoff structure in Equation (1) has a variety of applications [Ballester et al., 2006]. For example, Equation (1) can be interpreted as the profit function of a bank competing in quantities of lending à la Cournot with other banks in a single homogeneous product (a loan) on the same market (see Cohen-Cole et al. [2011]). Equation (1) can also be interpreted as the utility of a representative consumer in country i. Goods produced in different countries j trading with country i are complements, with a parameter $\lambda \geq 0$, and the utility of i has an own concavity term in domestically produced goods.

In order to find the Nash equilibrium solution associated with the above payoff function (1), we define a network centrality measure introduced by Bonacich [1987]. Let $\lambda_{\mathrm{PF}}(G)$ be the largest real eigenvalue of the adjacency matrix $\mathbf{A}$ of network G. The adjacency matrix is a matrix that lists the direct connections in the network. If $\mathbf{I}$ denotes the $n \times n$ identity matrix and $\mathbf{u} \equiv (1,\ldots,1)^\top$ the n-dimensional vector of ones then we can define *Bonacich centrality* as follows:

Definition 1. *If and only if $\lambda < 1/\lambda_{PF}(G)$ then the matrix $\mathbf{B}(G,\lambda) \equiv (\mathbf{I} - \lambda\mathbf{A})^{-1} = \sum_{k=0}^\infty \lambda^k \mathbf{A}^k$ exists, is non-negative,*[11] *and the vector of Bonacich centralities is defined as*

$$\mathbf{b}(G,\lambda) \equiv \mathbf{B}(G,\lambda) \cdot \mathbf{u}. \tag{2}$$

We can write the vector of Bonacich centralities as $\mathbf{b}(G,\lambda) = \sum_{k=0}^\infty \lambda^k \mathbf{A}^k \cdot \mathbf{u} = (\mathbf{I} - \lambda\mathbf{A})^{-1} \cdot \mathbf{u}$. For the components $b_i(G,\lambda)$, $i = 1,\ldots,n$, we get

$$b_i(G,\lambda) = \sum_{k=0}^\infty \lambda^k (\mathbf{A}^k \cdot \mathbf{u})_i = \sum_{k=0}^\infty \lambda^k \sum_{j=1}^n \left(\mathbf{A}^k\right)_{ij}, \tag{3}$$

where $\left(\mathbf{A}^k\right)_{ij}$ is the ij-th entry of $\mathbf{A}^k$. Because $\sum_{j=1}^n \left(\mathbf{A}^k\right)_{ij}$ is the number of all walks of length k in G starting from i, $b_i(G,\lambda)$ is the number of all walks in G starting from i, where the walks of length k are weighted by their geometrically decaying factor λ^k.

Now we can turn to the equilibrium analysis of the game.

Theorem 1 (Ballester et al. [2006])**.** *Let $\mathbf{b}(G,\lambda)$ be the Bonacich network centrality of parameter λ. For $\lambda < 1/\lambda_{PF}(G)$ the unique interior Nash equilibrium solution of the simultaneous n–player move game with payoffs given by Equation (1) and strategy space $\mathbb{R}_+^n$ is given by*

$$x_i^* = b_i(G,\lambda), \tag{4}$$

for all $i = 1,\ldots,n$.

[11] The proof can be found e.g. in Debreu and Herstein [1953].

Moreover, the payoff of agent i in the equilibrium is given by[12]

$$\pi_i^*(G,\lambda) \equiv \pi_i(\mathbf{x}^*, G, \lambda) = \frac{1}{2}(x_i^*)^2 = \frac{1}{2}b_i^2(G,\lambda). \tag{5}$$

The parameter λ measures the effect on agent i of agent j's contribution, if they are connected. If we assume that we have strong network externalities so that λ approaches its highest possible value $1/\lambda_{\mathrm{PF}}(G)$ then the Bonacich centrality becomes proportional to the standard eigenvector measure of centrality [Wasserman and Faust, 1994, Chap. 5.2]. The latter result has been shown by Bonacich [1987] and Bonacich and Lloyd [2001].

Furthermore, Ballester et al. [2006] have shown that the equilibrium outcome and the payoff for each player increases with the number of links in G (because the number of network walks increases in this way).[13] This implies that, if an agent is given the opportunity to change her links, she will add as many links as possible. On the other hand, if she is only allowed to form one link at a time, she will form the link to the agent that increases her payoff the most. In both cases, eventually, the network will then become complete, i.e. each agent is connected to every other agent. However, to avoid this latter unrealistic situation, we assume that the agents are living in a volatile environment that causes links to decay such that the complete network can never be reached. Instead the architecture of the network adapts to the volatile environment. We will treat these issues more formally in the next section.

3.2. The Network Formation Process

We now introduce a network formation process that incorporates the idea that agents with high Bonacich centrality (their equilibrium effort levels) are more likely to connect to each other, while the links they have established between each other have a longer life time if they are viewed as more valuable to them.

We consider continuous time Markov chain $(G(t))_{t\in\mathbb{R}_+}$ with $G(t) = (\mathcal{N}, \mathcal{E}(t))$ comprising the set of agents $\mathcal{N} = \{1, \ldots, n\}$ together with the set of edges/links $\mathcal{E}(t) \subset \mathcal{N} \times \mathcal{N}$ at time t between them. $(G(t))_{t\in\mathbb{R}_+}$ is a collection of random variables $G(t)$, indexed by time $t \in \mathbb{R}_+$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is the countable state space of all networks with n nodes, $\mathcal{F}$ is the σ-algebra $\sigma(\{G(t) : t \in \mathbb{R}_+\})$ generated by the collection of $G(t)$, and $\mathbb{P}\colon \mathcal{F} \to [0,1]$ is a countably additive, non-negative measure on $(\Omega, \mathcal{F})$ with total mass $\sum_{G\in\Omega} \mathbb{P}(G) = 1$. At every time $t \geq 0$, links can be created or decay with specified rates that depend on the current network $G(t) \in \Omega$.

Definition 2. *Consider a continuous time Markov chain $(G(t))_{t\in\mathbb{R}_+}$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\pi^*(G(t),\lambda) \equiv (\pi_1^*(G(t)), \ldots, \pi_n^*(G(t)))$ denote the profile of Nash equilibrium payoffs of the agents in $G(t)$ following from the payoff function (1) with parameter $0 \leq \lambda < 1/\lambda_{PF}(G(t))$.*[14]

[12] As discussed in more detail in Section 9, all our results would be unchanged in a setup where agents do not choose efforts but where their payoff is given by any monotonic increasing function of their Bonacich centrality.

[13] See Theorem 2 in Ballester et al. [2006].

[14] In order to guarantee an interior solution of the Nash equilibrium efforts corresponding to the payoff functions in Equation (1), we assume that the parameter $\lambda \geq 0$ is smaller than the inverse of the largest real eigenvalue of $G(t)$ for any t. Testing the impact of the Bonacich centrality measure on educational

(i) At a rate $\alpha_i \in (0,1)$ link creation opportunities arrive at each agent $i \in \mathcal{N}$. If such an opportunity arrives, then agent i computes the marginal payoff $\pi_i^*(G(t) \oplus ij, \lambda)$ for each agent $j \notin \mathcal{N} \backslash (\mathcal{N}_i \cup \{i\})$ she is not already connected to, where this computation includes an additive, exogenous stochastic term ε_{ij}, incorporating possible mistakes in the computation of the agent [cf. Blume, 2003; Hofbauer and Sandholm, 2007].[15] *We assume that the exogenous random terms ε_{ij} are identically and independently type I extreme value distributed (or Gumbel distributed) with scaling parameter ζ [cf. McFadden, 1981].*[16] *Given that agent i receives a link creation opportunity, she then links to agent j with probability [cf. Anderson et al., 1992]*[17]

$$\begin{aligned} b_i^\zeta(j|G(t)) &\equiv \mathbb{P}\left(\pi_i^*(G(t)\oplus ij,\lambda)+\varepsilon_{ij} = \max_{k\in\mathcal{N}\backslash(\mathcal{N}_i\cup\{i\})} \pi_i^*(G(t)\oplus ik,\lambda)+\varepsilon_{ik}\right)\mathbb{1}_{\mathcal{N}\backslash(\mathcal{N}_i\cup\{i\})}(j) \\ &= \frac{e^{\pi_i^*(G(t)\oplus ij,\lambda)/\zeta}}{\sum_{k\in\mathcal{N}\backslash(\mathcal{N}_i\cup\{i\})} e^{\pi_i^*(G(t)\oplus ik,\lambda)/\zeta}}\mathbb{1}_{\mathcal{N}\backslash(\mathcal{N}_i\cup\{i\})}(j). \end{aligned}$$

It follows that the probability that in a small time interval $[t, t+\Delta t)$ a transition takes place from $G(t)$ to $G(t) \oplus ij$ is given by $\mathbb{P}(G(t+\Delta t) = G \oplus ij | G(t) = G) = \alpha_i b_i^\zeta(j|G(t))\Delta t + o(\Delta t)$.[18]

(ii) We assume that a link ij, once established, has an exponentially distributed life time $\tau_{ij} \in \mathbb{R}_+$ with parameter $\nu_{ij}^\zeta(G(t)) \equiv \frac{1}{\mathbb{E}(\tau_{ij}|G(t))} = \beta_i f_{ij}(G(t))$, including an agent specific component $\beta_i \in (0,1)$, and a link specific component

$$f_{ij}^\zeta(G(t)) \equiv \frac{e^{\pi_i^*(G(t)\ominus ij,\lambda)/\zeta}}{\sum_{k\in\mathcal{N}_i} e^{\pi_i^*(G(t)\ominus ik,\lambda)/\zeta}}\mathbb{1}_{\mathcal{N}_i}(j).$$

The probability that in a small time interval $[t, t+\Delta t)$ a transition takes place from $G(t)$ to $G(t) \ominus ij$ is given by $\mathbb{P}(G(t+\Delta t) = G \ominus ij | G(t) = G) = \beta_i f_{ij}^\zeta(G(t))\Delta t + o(\Delta t)$.

Transitions to networks that differ by more than one link have probability of the order $o(\Delta t)$.

In words, if agent i is chosen to form a link (at rate α_i), she will choose the agent that increases the most her utility. There is, however, a possibility of error, captured by the stochastic term in the profit function. Furthermore, it is assumed that links do not last forever but have an exponentially distributed life time with an expectation that depends on the relative payoff loss from removing that link. The specific functional form of the pairwise component $f_{ij}^\zeta(\cdot)$ in the expected life time of a link incorporates the fact that links which are more valuable to an agent (i.e. the ones with the highest Bonacich centrality) live longer

outcomes in the United States, Calvó-Armengol et al. [2009] found that only 18 out of 199 networks (i.e. 9 %) do not satisfy this eigenvalue condition. In Section 9, we show that our results still hold in a more general framework (i.e. a more general utility function) where this condition on the eigenvalue is not needed.

[15] See also Sandholm [2010, Chap. 6].

[16] For the distribution of the error term it holds that $\mathbb{P}(\varepsilon_{ij} \leq c) = e^{-e^{c/\zeta-\gamma}}$, where $\gamma \approx 0.58$ is Euler's constant. The expectation is $\mathbb{E}(\varepsilon_{ij}) = 0$ and the variance is given by $\mathrm{Var}(\varepsilon_{ij}) = \frac{\pi^2\zeta^2}{6}$.

[17] These are the *perturbed best response* strategies analyzed in Hofbauer and Sandholm [2007].

[18] $f(t) = o(g(t))$ as $t \to \infty$ if $\lim_{t\to\infty} f(t)/g(t) = 0$.

than the ones which are viewed as less valuable to her. The value of a link is measured by the perceived loss in payoff incurred by the agent from removing the link.[19] The fact that links do not last forever is a quite natural feature of real-world networks. Ehrhardt et al. [2008] put forward inter-firm alliances and scientific collaborations as examples of networks with volatile environment. In the context of inter-firm alliances, Hagedoorn [2002] for research partnerships, Kogut et al. [2007] for joint ventures, Harrigan [1988] for alliances and Park and Russo [1996] for (equity-based) joint ventures provide empirical evidence on this phenomenon. For example, Harrigan [1988] studies 895 alliances from 1924 to 1985 and concludes that the average life-span of the alliance is relatively short, 3.5 years, with a standard deviation of 5.8 years and 85 % of these alliances last less than 10 years. Park and Russo [1996] focus on 204 joint ventures among firms in the electronic industry for the period 1979-1988. They show that less than half of these firms remain active beyond a period of five years and for those that last less than 10 years (2/3 of the total), the average lifetime turns out to be 3.9 years. Similar empirical evidence can be found in Newman [2004], Grossman [2002], Goyal et al. [2006] for scientific collaborations in physics, biomedical research, computer science, mathematics and economics, showing a relative short life of link duration.

It should be clear that, when a new link may be added to the network, then that link proposal will always be accepted by the receiver. This is because it always increases the utility of the receiver due to local complementarities in the utility function. In fact, we will show below, that it will also be the best reply for the receiver (i.e. the best alternative in terms of link formation).

Observe that when agents decide to create a link, they do it in a *myopic* way, that is they only look at the agents that give them the *current* highest payoff. There is literature on farsighted networks where agents calculate their lifetime-expected utility when they want to create a link [see, e.g. Konishi and Ray, 2003]. We adopt here a myopic approach because of its tractability and because our model also incorporates effort decision.[20]

The Markov chain $(G(t))_{t\in\mathbb{R}_+}$ can be described infinitesimally in time by the *generator matrix* $\mathbf{Q}^\zeta$ with elements given by the *transition rates* $q^\zeta : \Omega \times \Omega \to \mathbb{R}$ defined by $\lim_{\Delta t\downarrow 0} \mathbb{P}(G(t+\Delta t) = G'|G(t) = G) = q^\zeta(G, G')$. Consequently, $q^\zeta(G, G \oplus ij) = \alpha_i b_i^\zeta(j|G)$ and $q^\zeta(G, G \ominus ij) = \beta_i f_{ij}^\zeta(G)$. The transition rates have the property that $q^\zeta(G, G') = q^\zeta(G, G \pm ij) \geq 0$ if G' differs from G by the link ij and $q^\zeta(G, G') = 0$ if G' differs from G by more than one link. Moreover, it must hold that $\sum_{G'\in\Omega} q^\zeta(G, G') = 0$, and one can show that $\mathbb{P}(G(t) = G'|G(0) = G) = e^{\mathbf{Q}^\zeta t}$. If a non-negative solution to $\mu^\zeta \mathbf{Q}^\zeta = 0$ with $\sum_{G\in\Omega} \mu^\zeta(G) = 1$ exists, then μ^ζ is the stationary distribution of the Markov chain satisfying $\mu^\zeta(G') = \lim_{t\to\infty} \mathbb{P}(G(t) = G'|G(0) = G)$ [see e.g. Ethier and Kurtz, 1986; Liggett, 2010; Stroock, 2005].

The most simple case is the one where ζ diverges, the error term ε_{ij} becomes dominant and the link formation and decay rates are payoff independent. The link creation and

[19] In a similar way, Staudigl [2011] assumes that the linking activity levels of agents depend on their relative marginal payoffs, and Snijders [2001]; Snijders et al. [2010] introduce exponential link update rates, which "depend on actor-specific covariates or on network statistics expressing the degree to which the actor is satisfied with the present network structure." See also Eq. (3.4) in Staudigl [2011] and Snijders [2001], Section 7.1.

[20] Jackson and Watts [2002b] argue that this form of myopic behavior makes sense if players discount heavily the future. In Section 9, we discuss the possibility of having farsighted agents in our model

decay rates are then given by

$$\lambda_i \equiv \lim_{\zeta\to\infty} q^{\zeta}(G, G\oplus ij) = \alpha_i \frac{1}{|\mathcal{N}\backslash(\mathcal{N}_i\cup\{i\})|}\mathbb{1}_{\mathcal{N}\backslash(\mathcal{N}_i\cup\{i\})}(j),$$
$$\mu_i \equiv \lim_{\zeta\to\infty} q^{\zeta}(G, G\ominus ij) = \beta_i \frac{1}{|\mathcal{N}_i|}\mathbb{1}_{\mathcal{N}_i}(j).$$

These transition rates correspond to a birth-death Markov chain with birth rates λ_i and death rates μ_i [see e.g. Liggett, 2010, Chap. 2.7.1]. It follows that the stationary degree distribution is given by $P(d) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{P}(d_i = d)$ where

$$\mathbb{P}(d_i = d) = \mathbb{P}(d_i = 0)\prod_{j=0}^{d-1}\frac{\lambda_j}{\mu_{j+1}} = \frac{1}{1+\frac{\alpha_i}{\beta_i}\sum_{d=1}^{n-1}\prod_{j=0}^{d-1}\frac{n-j}{j+2}}\frac{\alpha_i}{\beta_i}\prod_{j=0}^{d-1}\frac{n-j}{j+2}$$
$$= \frac{\alpha_i}{\alpha_i\left(2^{n+1}-n-3\right)+\beta_i(n+1)}\binom{n+1}{d+1}.$$

Setting $\alpha_i = \beta_i = 1/2$ for all $i \in \mathcal{N}$, we can write the degree distribution for large n as

$$P(d) = \frac{1}{2\left(2^n-1\right)}\binom{n+1}{d+1} \sim \left(\frac{1}{2}\right)^{n-1}\binom{n-1}{d}, \quad \text{as} \quad n\to\infty. \tag{6}$$

This is equivalent to a Poisson degree distribution $P(d) = \binom{n-1}{d}p^d(1-p)^{n-1-d}$ corresponding to a random graph $G(n,p)$ with an independent link probability $p = 1/2$ [see e.g. Jackson, 2008, Chap. 4].

A more interesting case, from a behavioral and topological point of view, is the one where ζ converges to zero and the error term ε_{ij} vanishes. For each agent $i \in \mathcal{N}$ let the best response be the set-valued map $\mathcal{B}_i : \Omega \to \mathcal{N}$ defined as $\mathcal{B}_i(G) \equiv \arg\max_{k\in\mathcal{N}\backslash(\mathcal{N}_i\cup\{i\})}\pi_i^*(G\oplus ik,\lambda)$, and similarly we define the map $\mathcal{M}_i : \Omega \to \mathcal{N}$ as $\mathcal{M}_i(G) \equiv \arg\max_{k\in\mathcal{N}_i}\pi_i^*(G\ominus ik,\lambda)$. In the limit $\zeta \to 0$, we then have that the link creation and decay rates are given by

$$q(G, G\oplus ij) \equiv \lim_{\zeta\to 0} q^{\zeta}(G, G\oplus ij) = \alpha_i\frac{1}{|\mathcal{B}_i(G)|}\mathbb{1}_{\mathcal{B}_i(G)}(j),$$
$$q(G, G\ominus ij) \equiv \lim_{\zeta\to 0} q^{\zeta}(G, G\ominus ij) = \beta_i\frac{1}{|\mathcal{M}_i(G)|}\mathbb{1}_{\mathcal{M}_i(G)}(j).$$

Networks in the support of the stationary distribution of the Markov chain when ζ converges to zero correspond to the *stochastically stable* states [Foster and Young, 1990].[21] A network $G \in \Omega$ is called stochastically stable if $\mu(G) > 0$ where $\mu \equiv \lim_{\zeta\to 0}\mu^{\zeta}$. The set of stochastically stable networks is denoted by $\hat{\Omega} \equiv \{G \in \Omega : \mu(G) > 0\}$. We will analyze these states in Section 4, while we will study the sample paths generated by the chain in the limit of ζ converging to zero in the next section.

[21] See also Young [1998, Chap. 3] and Sandholm [2010, Chap. 12].

3.3. Network Formation and Nested Split Graphs

In this section we will focus on the Markov chain introduced in Definition 2 as the probability of mistakes converges to zero. An essential property of the chain is that it produces networks in a well defined class of graphs denoted by "nested split graphs" [Aouchiche et al., 2008].[22] We will give a formal definition of these graphs and discuss an example in this section. Nested split graphs include many common networks such as the star or the complete network. Moreover, as their name already indicates, they have a *nested neighborhood structure*. This means that the set of neighbors of each agent is contained in the set of neighbors of each higher degree agent. Nested split graphs have particular topological properties and an associated adjacency matrix with a well defined structure.

In order to characterize nested split graphs, it will be necessary to consider the degree partition of a graph, which is defined as follows:

Definition 3 (Mahadev and Peled [1995]). *Let $G = (\mathcal{N}, \mathcal{E})$ be a graph whose distinct positive degrees are $d_{(1)} < d_{(2)} < \ldots < d_{(k)}$, and let $d_0 = 0$ (even if no agent with degree 0 exists in G). Further, define $\mathcal{D}_i = \{v \in \mathcal{N} : d_v = d_{(i)}\}$ for $i = 0, \ldots, k$. Then the set-valued vector $\mathcal{D} = (\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_k)$ is called the degree partition of G.*

With the definition of a degree partition, we can now give a more formal definition of a nested split graph.[23,24]

Definition 4 (Mahadev and Peled [1995]). *Consider a nested split graph $G = (\mathcal{N}, \mathcal{E})$ and let $\mathcal{D} = (\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_k)$ be its degree partition. Then the nodes $\mathcal{N}$ can be partitioned in independent sets $\mathcal{D}_i$, $i = 1, \ldots, \left\lfloor \frac{k}{2} \right\rfloor$ and a dominating set $\bigcup_{i=\left\lfloor \frac{k}{2} \right\rfloor+1}^{k} \mathcal{D}_i$ in the graph $G' = (\mathcal{N} \backslash \mathcal{D}_0, \mathcal{E})$. Moreover, the neighborhoods of the nodes are nested. In particular, for each node $v \in \mathcal{D}_i$, $i = 1, \ldots, k$,*

$$\mathcal{N}_v = \begin{cases} \bigcup_{j=1}^{i} \mathcal{D}_{k+1-j} & \text{if } i = 1, \ldots, \left\lfloor \frac{k}{2} \right\rfloor, \\ \bigcup_{j=1}^{i} \mathcal{D}_{k+1-j} \setminus \{v\} & \text{if } i = \left\lfloor \frac{k}{2} \right\rfloor + 1, \ldots, k. \end{cases} \tag{7}$$

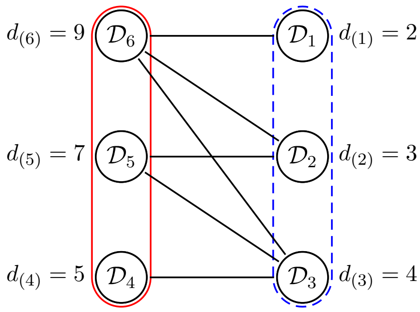
Figure 1 (left) illustrates the degree partition $\mathcal{D} = (\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_6)$ and the nested neighborhood structure of a nested split graph. A line between $\mathcal{D}_i$ and $\mathcal{D}_j$ indicates that every node in $\mathcal{D}_i$ is linked to every node in $\mathcal{D}_j$ for any $i, j = 1, \ldots, 6$. The nodes in the dominating set included in the solid frame induce a clique while the nodes in the independent sets that are included in the dashed frame induce an empty subgraph. In the following we will call the sets $\mathcal{D}_i$, $i = \left\lfloor \frac{k}{2} \right\rfloor + 1, \ldots, k$, *dominating subsets*, since the set $\mathcal{D}_i$ induces a dominating set in the graph obtained by removing the nodes in the set $\bigcup_{j=0}^{k-i} \mathcal{D}_j$ from G.

A nested split graph has an associated adjacency matrix which is called *stepwise matrix* and it is defined as follows:

[22] Nested split graphs are also called "threshold networks" [Hagberg et al., 2006; Mahadev and Peled, 1995].

[23] Let x be a real valued number $x \in \mathbb{R}$. Then, $\lceil x \rceil$ denotes the smallest integer larger or equal than x (the ceiling of x). Similarly, $\lfloor x \rfloor$ denotes the largest integer smaller or equal than x (the floor of x).

[24] In general, split graphs are graphs whose nodes can be partitioned in a set of nodes which are all connected among each other and sets of nodes which are disconnected. A nested split graph is a special case of a split graph.



$$\mathbf{A} = \begin{pmatrix} 0&1&1&1&1&1&1&1&1&1\\ 1&0&1&1&1&1&1&1&1&1\\ 1&1&0&1&1&1&1&1&0&0\\ 1&1&1&0&1&1&0&0&0&0\\ 1&1&1&1&0&0&0&0&0&0\\ 1&1&1&1&0&0&0&0&0&0\\ 1&1&1&0&0&0&0&0&0&0\\ 1&1&1&0&0&0&0&0&0&0\\ 1&1&0&0&0&0&0&0&0&0\\ 1&1&0&0&0&0&0&0&0&0 \end{pmatrix}$$

Figure 1: Representation of a connected nested split graph (left) and the associated adjacency matrix (right) with $n = 10$ agents and $k = 6$ distinct positive degrees. A line between $\mathcal{D}_i$ and $\mathcal{D}_j$ indicates that every node in $\mathcal{D}_i$ is linked to every node in $\mathcal{D}_j$. The solid frame indicates the dominating set and the nodes in the independent sets are included in the dashed frame. Next to the set $\mathcal{D}_i$ the degree of the nodes in the set is indicated. The neighborhoods are nested such that the degrees are given by $d_{(i+1)} = d_{(i)} + |\mathcal{D}_{k-i+1}|$ for $i \neq \lfloor \frac{k}{2} \rfloor$ and $d_{(i+1)} = d_{(i)} + |\mathcal{D}_{k-i+1}| - 1$ for $i = \lfloor \frac{k}{2} \rfloor$. In the corresponding adjacency matrix $\mathbf{A}$ to the right the zero-entries are separated from the one-entries by a stepfunction.

Definition 5 (Brualdi and Hoffman [1985])**.** *A stepwise matrix* $\mathbf{A}$ *is a matrix with elements* a_{ij} *satisfying the condition: if* $i < j$ *and* $a_{ij} = 1$ *then* $a_{hk} = 1$ *whenever* $h < k \leq j$ *and* $h \leq i$*.*

Figure 1 (right) shows the stepwise adjacency matrix $\mathbf{A}$ corresponding to the nested split graph shown on the left hand side. If we let the nodes by indexed by the order of the rows in the adjacency matrix $\mathbf{A}$ then it is easily seen that for example $\mathcal{D}_6 = \{1, 2 \in \mathcal{N} : d_1 = d_2 = d_{(6)} = 9\}$ and $\mathcal{D}_1 = \{9, 10 \in \mathcal{N} : d_9 = d_{10} = d_{(1)} = 2\}$.

If a nested split graph is connected we call it a connected nested split graph. The representation and the adjacency matrix depicted in Figure 1 actually shows a connected nested split graph. From the stepwise property of the adjacency matrix it follows that a connected nested split graph contains at least one spanning star, that is, there is at least one agent that is connected to all other agents. In Appendix B, we also derive the clustering coefficient, the neighbor connectivity and the characteristic path length of a nested split graph. In particular, we show that connected nested split graphs have small characteristic path length, which is at most two. We also analyze different measures of centrality (see Wasserman and Faust [1994, Chap. 5.2]) in a nested split graph. One important result is that degree, closeness, and Bonacich centrality induce the same ordering of nodes in a nested split graph. If the ordering is not strict, then this holds also for betweenness centrality (see Section B.2.5 in the Appendix).

In the next proposition, we identify the relationship between the Bonacich centrality of an agent and her degree in a nested split graph.

Proposition 1. *Consider a pair of agents* $i, j \in \mathcal{N}$ *of a nested split graph* $G = (\mathcal{N}, \mathcal{E})$*.*

(i) *If and only if agent* i *has a higher degree than agent* j *then* i *has a higher Bonacich centrality than* j*, i.e.* $d_i > d_j \Leftrightarrow b_i(G, \lambda) > b_j(G, \lambda)$*.*

(ii) *Assume that neither the links* ik *nor* ij *are in* G*,* $ij \notin \mathcal{E}$ *and* $ik \notin \mathcal{E}$*. Further assume that agent* k *has a higher degree than agent* j*,* $d_k > d_j$*. Then adding the link* ik *to* G *increases the Bonacich centrality of agent* i *more than adding the link* ij *to* G*, i.e.* $d_k > d_j \Leftrightarrow b_i(G \oplus ik, \lambda) > b_i(G \oplus ij, \lambda)$*.*

From part (ii) of Proposition 1 we find that when agent i has to decide to create a link either to agents k or j, with $d_k > d_j$, in the link formation process $(G(t))_{t\in\mathbb{R}_+}$ then i will always connect to agent k because this link gives i a higher Bonacich centrality than the other link to agent j. We can make use of this property in order to show that the networks emerging from the link formation process defined in the previous section actually are nested split graphs. This result is stated in the next proposition.

Proposition 2. *Consider the network formation process $(G(t))_{t\in\mathbb{R}_+}$ introduced in Definition 2 in the limit of ζ converging to zero. Assume that at $t = 0$, we start with the empty network $G(0) = \bar{K}_n$. Then, at any time $t \geq 0$, the network $G(t)$ is a nested split graph almost surely, and the set $\Psi \in \Omega$ consisting of all possible unlabeled nested split graphs on n nodes with $|\Psi| = 2^{n-1}$ has measure $\mathbb{P}(\Psi) = 1$.*

This result is due to the fact that agents, when they have the possibility of creating a new link, always connect to the agent who has the highest Bonacich centrality (and by Proposition 1 the highest degree). This creates a nested neighborhood structure which can always be represented by a stepwise adjacency matrix after a possible relabeling of the agents.[25] The same applies for link decay.

Let us give some more intuition of this crucial result. Agents want to link to others who are more central since this leads to higher actions (as actions are proportional to centrality) and higher actions raise payoffs more. Similarly, links decay to those with lower centrality as these agents have lower actions and hence lower payoff effects. Notice moreover that, once a link decays to an agent, she becomes less central and this makes it more likely that another link decays. Thus link gains and losses are self reinforcing. This intuition suggests that if α, the probability of adding links is large then the process should approximate complete network while if it is small then the process should approximate the star network. The key insight of our model is that for intermediate values of α the network is a nested split graph.

Observe that it is assumed that there is no cost to forming links. If links represent a social tie, then there is a cost to maintaining a link since agents must spend time with the person they are linked to. Because of the assumption of no link cost, each agent wants to connect to every other agent, which leads to the formation of nested split graphs. In Section 9 and Appendix C, we extend the model to see what would happen to the results if links were costly to maintain and only the links that increase the payoff of an agent were formed. We show that as long as the cost is not too high, marginal payoffs are positive and the networks always converges to nested split graphs so that all our results hold.

Due to the nested neighborhood structure of nested split graphs, any pair of agents in (the connected component) of a nested split graph is at most two links separated from each other. From Proposition 1 it then follows that in a nested split graph $G(t)$ the best response of an agent i are the agents with the highest degrees in i's second-order neighborhood $\mathcal{N}_i^{(2)}$.[26] Moreover, if $G(t)$ is a nested split graph then $i \in \mathcal{B}_j(G(t))$ if and

[25] Further, we will show in Proposition 3 that as ζ converges to zero, $(G(t))_{t\in\mathbb{R}_+}$ induces a finite state Markov chain where the recurrent states $\hat{\Omega}$ consist of nested split graphs.

[26] Let $\mathcal{N}_i = \{k \in \mathcal{N} : ik \in \mathcal{E}(t)\}$ be the set of neighbors of agent $i \in \mathcal{N}$ and $\mathcal{N}_i^{(2)} = \bigcup_{j\in\mathcal{N}_i} \mathcal{N}_j \backslash (\mathcal{N}_i \cup \{i\})$ denote the second-order neighbors of agent i in the current network $G(t)$. Note that the connectivity relation is symmetric such that j is a second-order neighbor of i if i is a second-order neighbor of j, i.e. $i \in \mathcal{N}_j^{(2)}$ if and only if $j \in \mathcal{N}_i^{(2)}$ for all $i, j \in \mathcal{N}$.

only if $j \in \mathcal{B}_i(G(t))$. Hence, we could require in addition that links are only formed under mutual consent as in the *consent game* by Myerson [1991].

From the fact that $G(t)$ is a nested split graph with an associated stepwise adjacency matrix it further follows that at any time t in the network evolution, $G(t)$ consists of a single connected component and possibly isolated nodes.

Corollary 1. *Consider the network formation process $(G(t))_{t\in\mathbb{R}_+}$ introduced in Definition 2 in the limit of ζ converging to zero. Assume that at $t = 0$, we start with the empty network $G(0) = \bar{K}_n$. Then, at any time $t \geq 0$, the network $G(t)$ has at most one non-singleton component almost surely.*

Nested split graphs are not only prominent in the literature on spectral graph theory [Cvetkovic et al., 1997] but they have also appeared in the recent literature on economic networks. Nested split graphs are so called "inter-linked stars" found in Goyal and Joshi [2003].[27] Subsequently, Goyal et al. [2006] identified inter-linked stars in the network of scientific collaborations among economists. It is important to note that nested split graphs are characterized by a distinctive core-periphery structure. Core-periphery structures have been found in several empirical studies of interfirm collaborations networks [Baker et al., 2008; Kitsak et al., 2010]. A recent study by Leskovec et al. [2009] also finds nested core-periphery structures in over 100 large sparse real-world social and information networks. The wider applicability of nested split graphs suggests that a network formation process as it is defined in Definition 2 that generates these graphs as the stochastically stable networks are of general relevance for understanding economic and social networks.

Finally, note that the network formation process $(G(t))_{t\in\mathbb{R}_+}$ introduced in Definition 2 is independent of initial conditions $G(0)$.[28] This means that even when we start from an initial network $G(0)$ which is not a nested split graph then after some finite time the Markov chain will reach a nested split graph. From then on all consecutive networks visited by the chain are nested split graphs.

4. Stochastically Stable Networks: Characterization

In this section we show that the network formation process $(G(t))_{t\in\mathbb{R}_+}$ of Definition 2 induces an ergodic Markov chain with a unique invariant distribution. We then proceed by analyzing the stochastically stables states in $\hat{\Omega}$ (in the limit of $\zeta \to 0$) of this process as the number n of agents becomes large.

Proposition 3. *The network formation process $(G(t))_{t\in\mathbb{R}_+}$ introduced in Definition 2 induces an ergodic Markov chain on the finite state space Ω with a unique stationary distribution μ^ζ such that $\mu^\zeta(G') = \lim_{t\to\infty} \mathbb{P}(G(t) = G'|G(0) = G)$ for any $G, G' \in \Omega$. Moreover, the stochastically stable states $\hat{\Omega}$ are given by the set of nested split graphs Ψ such that $\lim_{\zeta\to 0} \mu^\zeta(\Psi) = 1$.*

[27] Nested split graphs are inter-linked stars but an inter-linked star is not necessarily a nested split graph. Nested split graphs have a nested neighborhood structure for all degrees while in an inter-linked star this holds only for the nodes with the lowest and highest degrees.

[28] See Proposition 3 in Section 4 and its proof in Appendix D.

In the following we will assume for simplicity that $\alpha_i = 1 - \beta_i = \alpha$ for all $i \in \mathcal{N}$ in Definition 2, expressing the relative weights of link creation versus link decay.[29] In this case, the symmetry of the network formation process with respect to the link arrival rate α and the link decay parameter $1 - \alpha$ allows us to state the following proposition.

Proposition 4. *Consider the Markov chain* $(G(t))_{t\in\mathbb{R}_+}$ *of Definition 2 with* $\alpha \equiv \alpha_i = 1-\beta_i$ *for all* $i \in \mathcal{N}$. *Let* $G(t)$ *be a sample path generated with the homogeneous link arrival rate* α, *and let* $G'(t)$ *be a sample path with arrival rate* $1 - \alpha$. *Let* μ *be the stationary distribution of* $G(t)$ *and* μ' *the stationary distribution of* $G'(t)$ *in the limit* $\zeta \to 0$. *Then for each network* G *in the support of* μ *the complement* $\bar{G}$ *of* G *has the same probability in* μ', *i.e.* $\mu'(\bar{G}) = \mu(G)$.

Proposition 4 allows us to derive the stationary distribution μ for any value of $1/2 < \alpha < 1$ if we know the corresponding distribution for $1 - \alpha$. This follows from the fact that the complement $\bar{G}$ of a nested split graph G is a nested split graph as well [Mahadev and Peled, 1995]. In particular, the networks $\bar{G}$ are nested split graphs in which the number of nodes in the dominating subsets corresponds to the number of nodes in the independent sets in G and, conversely, the number of nodes in the independent sets in $\bar{G}$ corresponds to the number of nodes in the dominating subsets in G.

With this symmetry in mind we restrict our analysis in the following to the case of $0 < \alpha \leq 1/2$. Let $\{\mathbf{N}(t)\}_{t\in\mathbb{R}_+}$ be the degree distribution with the d-th element $N_d(t)$, giving the number of nodes with degree d in $G(t)$, in the t-th sequence $\mathbf{N}(t) \equiv \{N_d(t)\}_{d=0}^{n-1}$. Further, let $P_t(d) \equiv N_d(t)/n$ denote the proportion of nodes with degree d ($\mathbf{P}(t) \equiv \frac{1}{n}\mathbf{N}(t)$) and let $P(d) \equiv \lim_{t\to\infty} P_t(d)$ be its asymptotic value. In the following proposition we determine the asymptotic degree distribution of the nodes in the independent sets for n large enough.

Proposition 5. *Consider the Markov chain* $(G(t))_{t\in\mathbb{R}_+}$ *of Definition 2 with* $\alpha \equiv \alpha_i = 1-\beta_i$ *for all* $i \in \mathcal{N}$ *and let* $0 < \alpha \leq 1/2$. *Let* $P_t(d)$ *denote the proportion of nodes with degree* d *in* $G(t)$. *Then the asymptotic expected proportion of nodes in the independent sets with degrees,* $d = 0, 1, \ldots, d^*$, *in the stochastically stable networks* $G \in \hat{\Omega}$ *for large* n *is given by*

$$\lim_{t\to\infty} \mathbb{E}_t\left(P_t(d)\right) = \frac{1-2\alpha}{1-\alpha}\left(\frac{\alpha}{1-\alpha}\right)^d, \tag{8}$$

where[30]

$$d^*(n,\alpha) = \frac{\ln\left(\frac{(1-2\alpha)n}{2(1-\alpha)}\right)}{\ln\left(\frac{1-\alpha}{\alpha}\right)}, \tag{9}$$

and $P_t(d) \to \mathbb{E}_t\left(P_t(d)\right)$ *almost surely as* $n \to \infty$.

The structure of nested split graphs implies that if there exist nodes for all degrees between 0 and d^* (in the independent sets), then the dominating subsets with degrees

[29] Note that taking into account the possibility of an agent remaining quiescent only modifies the time-scale of the process discussed, thus yielding identical results to the model proposed. This implies that, without any loss of generality, it is possible to assume $\alpha_i + \beta_i = 1$ for all $i \in \mathcal{N}$. For simplicity, we also assume that these probabilities are the same across agents.

[30] Note that $d^*(n,\alpha)$ from Equation (9) might in general not be an integer. In this case we take the closest integer value to Equation (9), that is, we take $[d^*(n,\alpha)] = \lfloor d^*(n,\alpha) + \frac{1}{2} \rfloor$. The error we make in this approximation is negligible for large n.

larger than d^* contain only a single node. Further, using Proposition 4, we know that for $\alpha > 1/2$ the expected number of nodes in the dominating subsets is given by the expected number of nodes in the independent sets in Equation (8) for $1-\alpha$, while each of the independent sets contains a single node. This determines the asymptotic degree distribution for the independent or dominating subsets, respectively, for all values of α in the limit of large n.

From Equation (9) we can directly derive the following corollary.

Corollary 2. *Consider the Markov chain $(G(t))_{t\in\mathbb{R}_+}$ of Definition 2 with $\alpha_i = 1-\beta_i = \alpha$ for all $i \in \mathcal{N}$. Then there exists a phase transition in the asymptotic average number of independent sets, $d^*(n,\alpha)$, for $G \in \hat{\Omega}$ as n becomes large such that*

$$\lim_{n\to\infty} \frac{d^*(n,\alpha)}{n} = \begin{cases} 0, & \text{if } \alpha < \frac{1}{2}, \\ \frac{1}{2}, & \text{if } \alpha = \frac{1}{2}, \\ 1, & \text{if } \alpha > \frac{1}{2}. \end{cases} \tag{10}$$

Corollary 2 implies that as n grows without bound the networks in the stationary distribution μ are either sparse or dense, depending on the value of the link creation probability α. Moreover, from the functional form of $d(n,\alpha)$ in Equation (9) we find that there exists a sharp transition from sparse to dense networks as α crosses $1/2$ and the transition becomes sharper the larger is n.

Observe that, because a nested split graph is uniquely defined by its degree distribution,[31] Proposition 5 delivers us a complete description of a typical network generated by our model in the limit of large t and n. We call this network the "stationary network". We can compute the degree distribution and the corresponding adjacency matrix of the stationary network for different values of α.[32] The latter is shown in Figure 2. From the structure of these matrices we observe the transition from sparse networks containing a hub and many agents with small degree to a quite homogeneous network with many agents having similar high degrees. Moreover, this transition is sharp around $\alpha = 1/2$. In Figure 3, we show particular networks arising from the network formation process for the same values of α. Again, we can identify the sharp transition from hub-like networks to homogeneous, almost complete networks.

Figure 4 (left) displays the number $\bar{m}$ of links m relative to the total number of possible links $n(n-1)/2$, i.e. $\bar{m} = \frac{2m}{n(n-1)}$, and the number of distinct degrees k as a function of α. We see that there exists a sharp transition from sparse to dense networks around $\alpha = 1/2$ while k reaches a maximum at $\alpha = 1/2$. This follows from the fact that $k = 2d^*$ with d^* given in Equation (9) is monotonic increasing in α for $\alpha < 1/2$ and monotonic decreasing in α for $\alpha > 1/2$.

[31] The degree distribution uniquely determines the corresponding nested split graph up to a permutation of the indices of nodes.

[32] Non-integer values for the partition sizes can be approximated with the closest integer while preserving the nested structure of the degree partitions.

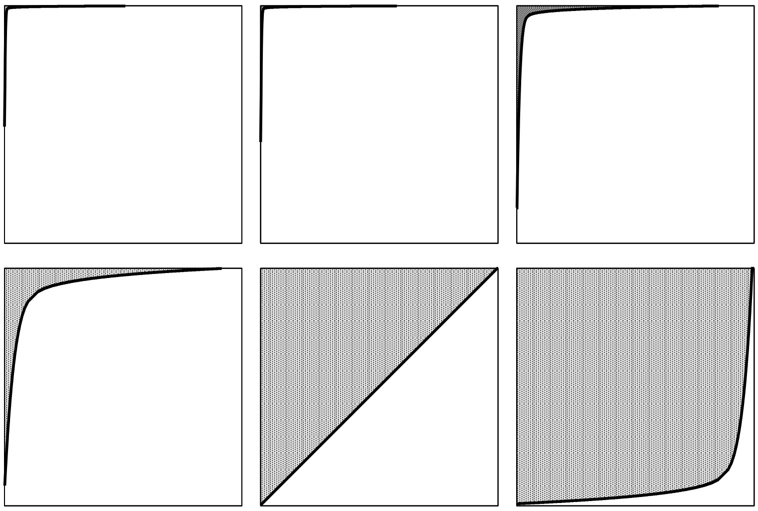


Figure 2: Representation of the adjacency matrices of stationary networks with $n = 1000$ agents for different values of parameter α: $\alpha = 0.2$ (top-left plot), $\alpha = 0.4$ (top-center plot), $\alpha = 0.48$ (top-right plot), $\alpha = 0.495$ (bottom-left plot), $\alpha = 0.5$ (bottom-center plot), and $\alpha = 0.52$ (bottom-right plot). The solid line illustrates the stepfunction separating the zero from the one entries in the matrix. The matrix top-left for $\alpha = 0.4$ is corresponding to a star-like network while the matrix bottom-right for $\alpha = 0.52$ corresponds to an almost complete network. Thus, there exists a sharp transition from sparse to densely connected stationary networks around $\alpha = 0.5$. Networks of smaller size for the same values of α can be seen in Figure 3.

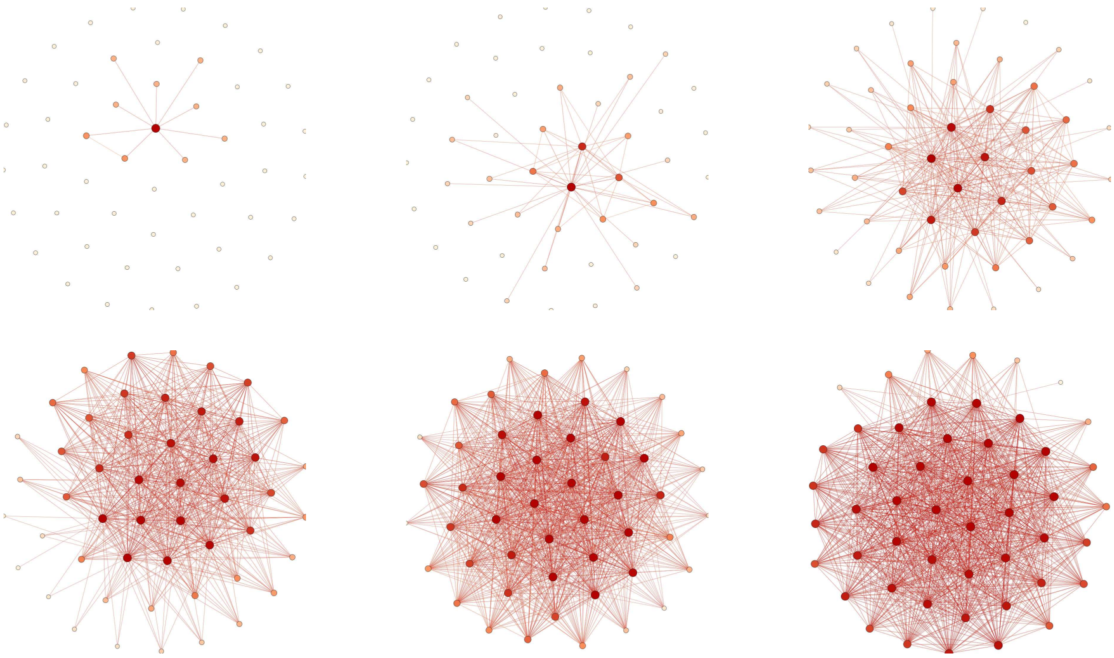


Figure 3: Sample networks with $n = 50$ agents for different values of parameter α: $\alpha = 0.2$ (top-left plot), $\alpha = 0.4$ (top-center plot), $\alpha = 0.48$ (top-right plot), $\alpha = 0.495$ (bottom-left plot), $\alpha = 0.5$ (bottom-center plot), and $\alpha = 0.52$ (bottom-right plot). The shade and size of the nodes indicate their eigenvector centrality. The networks for small values of α are characterized by the presence of a hub and a growing cluster attached to the hub. With increasing values of α the density of the network increases until the network becomes almost complete.

5. Stochastically Stable Networks: Statistics

We would like now to investigate further the properties of our networks and see how they match real-world networks. There exists a growing number of empirical studies trying to identify the key characteristics of social and economic networks. However, only few theoretical models (a notable exception is Jackson and Rogers [2007]) have tried to reproduce these findings to the full extent. We pursue the same approach. We show that our network formation model leads to properties which are shared with empirical networks. These properties can be summarized as follows:[33]

(i) The average shortest path length between pairs of agents is small [Albert and Barabási, 2002].

(ii) Empirical networks exhibit high clustering [Watts and Strogatz, 1998]. This means that the neighbors of an agent are likely to be connected.

(iii) The distribution of degrees is highly skewed. While some authors [e.g. Barabasi and Albert, 1999] find power-law degree distributions, others find deviations from power-laws in empirical networks, e.g. in Newman [2004], or exponential distributions [Guimera et al., 2006].

(iv) Several authors have found that there exists an inverse relationship between the clustering coefficient of an agent and her degree [Goyal et al., 2006; Pastor-Satorras et al., 2001]. The neighbors of a high degree agent are less likely to be connected among each other than the neighbors of an agent with low degree. This means that empirical networks are characterized by a negative clustering-degree correlation.

(v) Networks in economic and social contexts exhibit degree-degree correlations. Newman [2002, 2003] has shown that many social networks tend to be positively correlated. In this case the network is said to be assortative. On the other hand, technological networks such as the internet [Pastor-Satorras et al., 2001] display negative correlations. In this case the network is said to be dissortative. Others, however, find also negative correlations in social networks such as in the Ham radio network of interactions between amateur radio operators [Killworth and Bernard, 1976] or the affiliation network in a Karate club [Zachary, 1977]. Networks in economic contexts may have features of both technological and social relationships [Jackson, 2008] and so there exist examples with positive degree correlations such as in the network between venture capitalists [Mas et al., 2007] as well as negative degree correlations as it can be found in the world trade web [Serrano and Boguñá, 2003], online social communities [Hu and Wang, 2009] and in networks of banks [De Masi and Gallegati, 2007; May et al., 2008].

In the following sections, we analyze some of the topological properties of the stochastically stable networks in our model that are in the support of the stationary distribution μ. We simply refer to these networks as stationary networks. With the asymptotic expected degree distribution derived in Proposition 5, we can calculate the expected clustering coefficient, the clustering-degree correlation, the neighbor connectivity, the assortativity, and the characteristic path length by using the expressions derived for these quantities in Ap-

[33]This list of empirical regularities is far from being extensive and summarizes only the most pervasive patterns found in the literature.

pendix B, where we show that these statistics are all functions of the degree distribution.[34] These network measures are interesting because they can be compared to key empirical findings of social and economic networks. In fact, we show that the stationary networks exhibit all the well-known stylized facts of real-world networks. Moreover, in König et al. [2010], we show that, by introducing capacity constraints in the number of links an agent can maintain, we are able to produce both, assortative as well as dissortative networks.

Note that since the stationary distribution μ is unique, we can recover the expected value of any statistic by averaging over a large enough sample of empirical networks generated by numerical simulations. We then superimpose the analytical predictions of the statistic derived from Proposition 5 with the sample averages in order to compare the validity of our theoretical results, also for small network sizes n. As we will show, there is a good agreement of the theory with the empirical results for all n.

Degree Distribution From Proposition 5, we find that the degree distribution follows an exponential decay with a power-law tail.[35] The power-law tail has an exponent of minus one.[36] Degree distributions with power-law tails have been found in empirical networks, e.g. in scientific collaboration networks Newman [2004]. For $\alpha = 1/2$ the degree distribution is uniform while for larger values of α most of the agents have a degree close to the maximum degree.

Clustering The clustering coefficient is shown in Figure 5 (left). We find that for practically all values of α, the clustering in the stationary networks is high. This finding is in agreement with the vast literature on social networks that have reported high clustering being a distinctive feature of social networks. Moreover, Goyal et al. [2006] have shown that there exists a negative correlation between the clustering coefficient of an agent and her degree. We find this property in the stationary networks as well, as it is shown in Figure 5 (right).

Assortativity and Nearest Neighbor Connectivity We now turn to the study of correlations between the degrees of the agents and their neighbors. This property is usually

[34]Any network statistic $f : \Omega \to \mathbb{R}$ we consider can be expressed as a function of the (empirical) degree distribution $\mathbf{P}_t : \Omega \to [0,1]^n$. Hence, we can compute the expectation as $\mathbb{E}_t(f) = \sum_{\mathbf{k}\in(0,\ldots,n)^n} f(\mathbf{k}/n)\mathbb{P}_t(\mathbf{P}_t = \mathbf{k}/n)$. In Proposition 5 we show that the degree distribution converges to its expected value with probability one. Therefore, we have that $\mathbb{E}_t(f) = \sum_{\mathbf{k}\in(0,\ldots,n)^n} f(\mathbf{k}/n)\mathbb{1}_{\mathbb{E}_t(\mathbf{P}_t)}(\mathbf{k}/n) = f(\mathbb{E}_t(\mathbf{P}_t))$ as $n \to \infty$.

[35]For $0 < \alpha \leq 1/2$ and n large enough the asymptotic expected degree distribution for the degrees d smaller or equal than d^* is given by an exponential function $P(d) = \frac{1-2\alpha}{1-\alpha}\exp\left(-\ln\left(\frac{1-\alpha}{\alpha}\right)d\right)$. On the other hand, if we assume (i) that the degree of a node in a dominating subset is symmetrically distributed around its expected value, (ii) we compute the integral over the probability density function by a rectangle approximation and (iii) further assume that the degree distribution obtained in this way has the same functional form for all degrees d larger than d^* then one can show that for $0 < \alpha \leq 1/2$ and n large enough the asymptotic expected degree distribution $P(d)$ is given by $P(d) = \frac{\alpha}{(1-2\alpha)n}d^{-1}$. The power-law tail of the degree distribution can be confirmed by the empirical distribution from a logarithmic binning of numerical simulations, as can be seen in Figure 4 (right).

[36]We can extend our model to obtain a degree distribution with an arbitrary power-law tail by making the probability of selecting an agent depending on the number of links she already has, while preserving the nested structure of the network she is embedded in. This extension is further discussed in König and Tessone [2011].

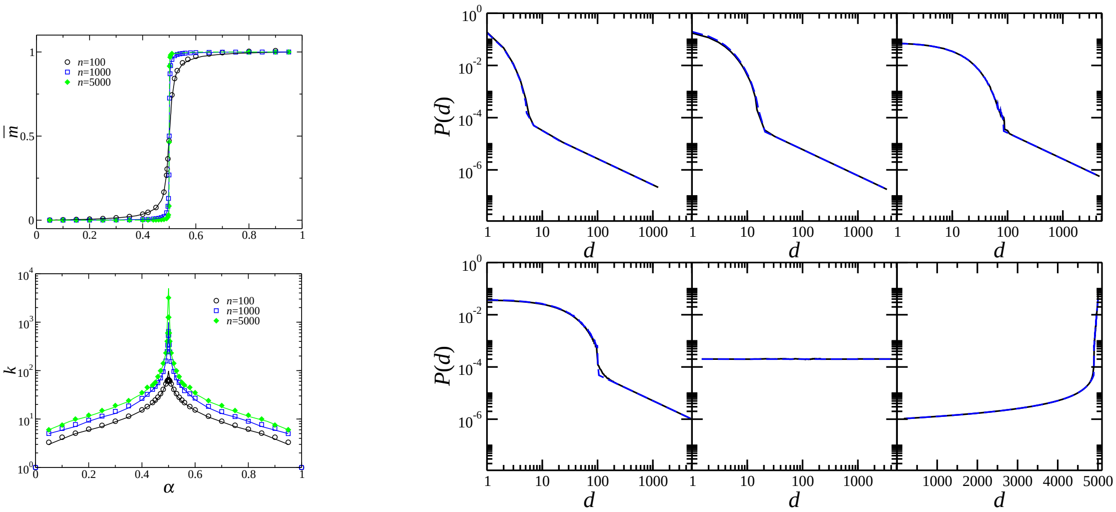


Figure 4: (Left) In the top panel we show the number $\bar{m}$ of links m relative to the total number of possible links $n(n-1)/2$ of the stationary network. The number of distinct degrees $k = 2d^*$, with d^* from Equation (9), found in the stationary network for different values of α are shown in the bottom panel. The figures display both, the results obtained by recourse of numerical simulations (symbols) and respecting theoretical predictions (lines) of the model. (Right) Degree distribution $P(d)$ for different values of parameter α and a network size $n = 10\,000$: $\alpha = 0.2$ (top-left plot), $\alpha = 0.4$ (top-center plot), $\alpha = 0.48$ (top-right plot), $\alpha = 0.49$ (bottom-left plot), $\alpha = 0.5$ (bottom-center plot), and $\alpha = 0.52$ (bottom-right plot). The solid line corresponds to the average of simulations while the dashed line indicates the theoretical degree distribution from Proposition 5. The degrees have been binned to smoothen the degree distribution.

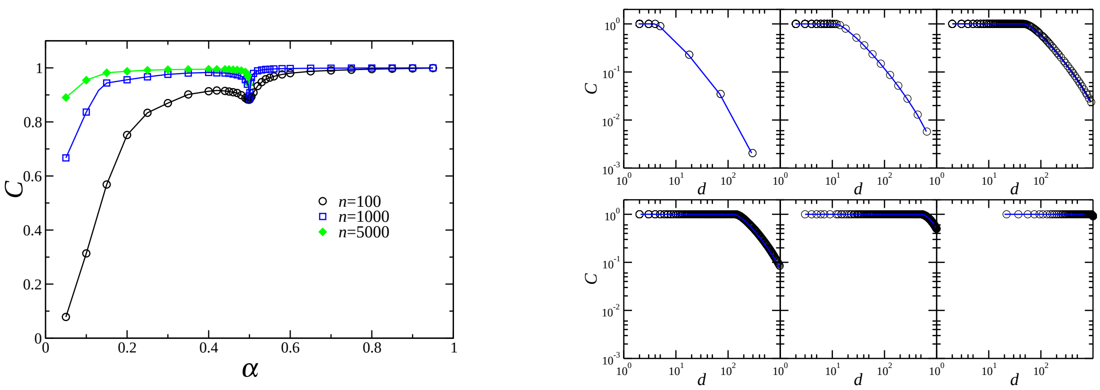


Figure 5: The left panel shows the clustering coefficient C and the right panel the clustering-degree correlation of stationary networks. The symbols correspond to the results obtained by recourse of numerical simulations. The solid lines correspond to the analytical results. We show that the clustering-degree correlation is negative for different values of α and a network size of $n = 1000$. The different plots show different values of α: $\alpha = 0.2$ (top-left plot), $\alpha = 0.4$ (top-center plot), $\alpha = 0.48$ (top-right plot), $\alpha = 0.49$ (bottom-left plot), $\alpha = 0.5$ (bottom-center plot), and $\alpha = 0.52$ (bottom-right plot).

measured by the network assortativity γ [Newman, 2002][37] and nearest neighbor connectivity $d_{nn}(d)$ [Pastor-Satorras et al., 2001]. Dissortative networks are characterized by negative degree correlations between a node and its neighbors and assortative networks show positive degree correlations. In dissortative networks γ is negative and $d_{nn}(d)$ monotonic decreasing while in assortative networks γ is positive and $d_{nn}(d)$ monotonic increasing. We find that in our basic model without capacity constraints (see König et al. [2010] for an extension including capacity constraints in the number of links an agent can maintain), we observe dissortative networks.

Assortativity and average nearest neighbor connectivity for different values of the link creation probability α are shown in Figure 6. Clearly, stationary networks are dissortative while the degree of dissortativity decreases with increasing α. However, if we recall the structure of the nested split graphs in Definition 4, to the class the stationary networks belong to, we can see that high degree agents are connected among each other while it is only the low degree agents that are not connected among each other. In this sense agents with high degrees tend to be connected to other agents with high degree. Considering only the agents with high degrees, we can call the network assortative. However, the agents with low degrees, that are only connected to agents with high degrees but are disconnected to agents with low degree, are so numerous in the stationary network (for low values of α) that we obtain an overall negative value for the assortativity of the network.

The dissortativity of stationary networks simply reflects the fact that stationary networks are strongly centralized for values of α below $1/2$. As an example consider a star $K_{1,n-1}$. $K_{1,n-1}$ is completely dissortative with an assortativity coefficient of $\gamma = -1$. Peripheral agents all have minimum degree one and are only connected to the central agent with maximum degree while the central agent is only connected to the agents with minimum degree. In this sense dissortativity is simply a measure of centralization in the network.

Characteristic Path Length Figure 8 shows the characteristic path length ℓ and the network efficiency ϵ (defined in Section B.1.4 in Appendix B). From these figures one can see that the characteristic path length ℓ never exceeds a distance of two. This means that for all parameter values of α stationary networks are characterized by short distances between agents. Together with the high clustering shown in this section the stationary networks can be seen as "small worlds" [Watts and Strogatz, 1998]. Stationary networks are efficient for values of α larger than $1/2$, in terms of short average distance between agents, while for values of α smaller than $1/2$ they are not. However, this short average distance is attained at the expense of a large number of links.

Centralization of Stochastically Stable Networks In the following section, we analyze the degree of centralization in stationary networks. As we will show, there exists a sharp transition in the centralization as a function of the link creation probability α. This means that stationary networks are either strongly centralized and hierarchical or

[37] The assortativity coefficient $\gamma \in [-1, 1]$ is essentially the Pearson correlation coefficient of degree between nodes that are connected. Positive values of γ indicate that nodes with similar degrees tend to be connected (and $d_{nn}(d)$ is an increasing function of the degree d) while negative values indicate that nodes with different degrees tend to be connected (and $d_{nn}(d)$ is a decreasing function of the degree d). See Newman [2002] for further details.

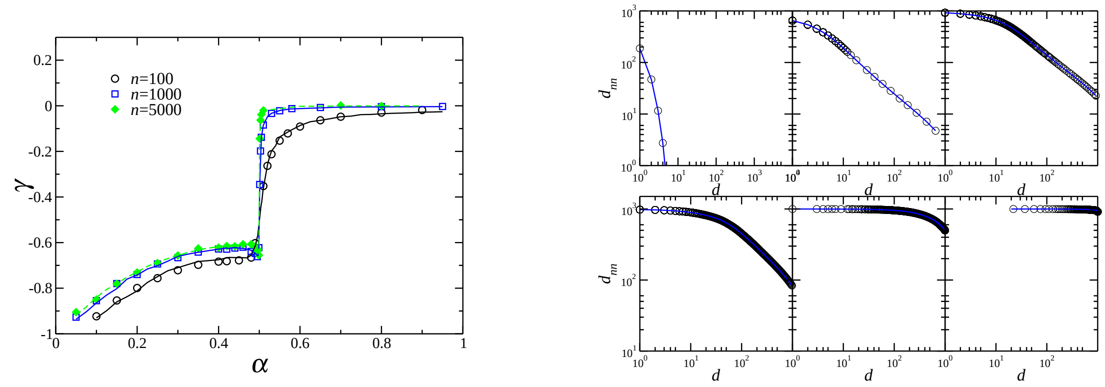


Figure 6: In the left panel we show the assortativity γ of stationary networks. In the right panel we show the average nearest neighbor connectivity d_{nn} for $\alpha = 0.2$ (top-left plot), $\alpha = 0.4$ (top-center plot), $\alpha = 0.48$ (top-right plot), $\alpha = 0.49$ (bottom-left plot), $\alpha = 0.5$ (bottom-center plot), and $\alpha = 0.52$ (bottom-right plot). The symbols correspond to the results obtained by recourse of numerical simulations. The solid lines correspond to the analytical results.

decentralized and homogeneous, depending on α. In Section 6, we will also find such a transition in the aggregate payoffs and effort levels of the agents.

For our analysis, we use the centralization index introduced by Freeman [1979]. The centralization $C : \Omega \to [0, 1]$ of a network $G = (\mathcal{N}, \mathcal{E}) \in \Omega$ is given by

$$C \equiv \frac{\sum_{u \in G} (C(u^*) - C(u))}{\max_{G'} \sum_{v \in G'} (C(v^*) - C(v))}, \tag{11}$$

where u^* and v^* are the agents with the highest values of centrality in the current network and and the maximum in the denominator is computed over all networks $G' = (\mathcal{N}, \mathcal{E}') \in \Omega$ with the same number of agents. For the degree, closeness, betweenness and eigenvector centrality measures one obtains the following indices[38]

$$C_d \equiv \frac{\sum_{u \in \mathcal{N}} (C_d(u^*) - C_d(u))}{n^2 - 3n + 2}, \qquad C_c \equiv \frac{\sum_{u \in \mathcal{N}} (C_c(u^*) - C_c(u))}{(n^2 - 3n + 2)/(2n - 3)},$$

$$C_b \equiv \frac{\sum_{u \in \mathcal{N}} (C_b(u^*) - C_b(u))}{n^3 - 4n^2 + 5n - 2}, \qquad C_v \equiv \frac{\sum_{u \in \mathcal{N}} (C_v(u^*) - C_v(u))}{\sqrt{(n-1)/2}(\sqrt{n-1} - 1)}.$$

From Figure 7 (right), showing degree, closeness, betweenness and eigenvector centralization, we clearly see that there exists a phase transition at $\alpha = 1/2$ from highly centralized to highly decentralized networks. This means that for low arrival rates of linking opportunities α (and a strong link decay) the stationary network is strongly polarized, composed mainly of a star (or an inter-linked star as in Goyal and Joshi [2003]), while for high arrival rates of linking opportunities (and a weak link decay) stationary networks are largely homogeneous. We can also see that the transition between these states is sharp. It is interesting to note that the same pattern emerges for all centrality measures considered,

[38]For the normalization of all the centralization indices we have used the star $K_{1,n-1}$. For degree, closeness, betweenness and eigenvector centralization it can be shown that $K_{1,n-1}$ is the network that maximizes the sum of differences in centrality [Bolland, 1988; Freeman, 1979].

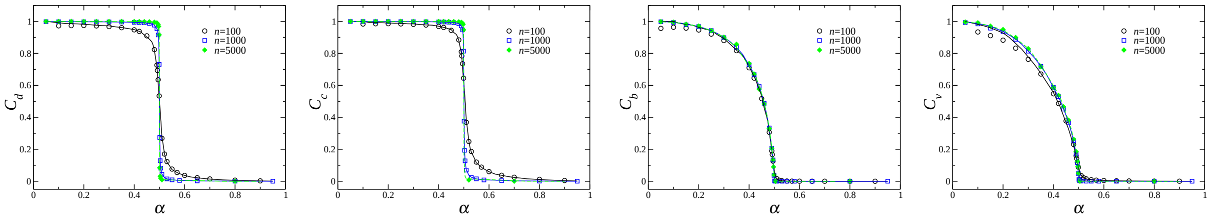


Figure 7: (From left to right) Degree, closeness, betweenness and eigenvector centralization in the stationary networks for different values of α. For all centralization measures we obtain a sharp transition between strongly centralized networks for lower values of α and decentralized networks for higher values of α. Note that we have only considered the connected component for the computation of the different centralization measures.

irrespective of whether the measures takes into account only the local neighborhood of an agent, such as in the case of degree centrality, or the entire network structure, as for the other centrality measures.[39]

Our findings are in line with previous works studying the optimal internal communication structure of organizations [Guimerà et al., 2002]. Other works [Calvó-Armengol et al., 2009; Dodds et al., 2003; Dupouet and Yildizoglu, 2006; Huberman and Hogg, 1995] have discussed the conditions under which informal organizational networks outperform centralized structures in complex, changing environments and under which conditions hierarchies are more efficient. Similar to Arenas et al. [2010] and Ehrhardt et al. [2006], we find sharp transitions between largely homogeneous and centralized networks. Moreover, the stationary networks in our model are polarized and strongly centralized for a low volatility in the environment associated with many linking opportunities whereas they are homogeneous and largely decentralized for a highly volatile environment with few linking opportunities and a strong link decay. The hierarchical structure of stationary networks and its dependency on the volatility is similar to the findings for optimal networks in Arenas et al. [2010].

6. Stochastically Stable Networks: Efficiency

We now turn to the investigation of the optimality and efficiency of stochastically stable networks. Following Jackson and Wolinsky [1996] and Jackson [2008], we define the social welfare as the sum of the agents' individual payoffs $\Pi(\mathbf{x}^*, G\lambda) = \sum_{i=1}^{n} \pi_i(\mathbf{x}^*, G, \lambda)$. We are interested in the solution of the following social planner's problem. Let $\mathcal{G}(n)$ denote the set of connected graphs having n agents in total. The social planner's solution is given by $G^* = \text{argmax}_{G \in \mathcal{G}(n)} \quad \Pi(\mathbf{x}^*, G, \lambda)$. A graph G^* solving the maximization problem will be denoted as "efficient". The efficient network has been derived in Ballester et al. [2006] and we state their result in the following proposition.

Proposition 6 (Ballester et al. [2006])**.** *Let $\mathcal{G}(n)$ denote the set of connected graphs having n agents and consider $G \in \mathcal{G}(n)$. Then the efficient network $G^* = \text{argmax}_{G \in \mathcal{G}(n)} \quad \Pi(\mathbf{x}^*, G, \lambda)$ is the complete graph K_n.*

[39]It is also irrelevant whether we use a centrality measure based on shortest paths, as for closeness and betweenness centrality, or one that is based on all paths, such as eigenvector centrality.

This proposition is a direct consequence of Theorem 2 in Ballester et al. [2006] where more links is always better. Moreover, Corbo et al. [2006] have shown that, in the case of strong complementarities, when λ approaches $1/\lambda_{\mathrm{PF}}(G)$, maximizing aggregate equilibrium payoffs is equivalent to maximizing the largest real eigenvalue $\lambda_{\mathrm{PF}}(G)$ of the network G.[40]

Proposition 7 (Corbo et al. [2006])**.** *Let $\mathcal{G}(n,m)$ denote the set of connected graphs having n agents and m links and consider $G \in \mathcal{G}(n,m)$. As $\lambda \uparrow 1/\lambda_{PF}(G)$, maximizing aggregate equilibrium contribution and payoff reduces to* $\mathrm{argmax}\{\Pi(\mathbf{x}^*, G, \lambda) : G \in \mathcal{G}(n,m)\} = \mathrm{argmax}\{\lambda_{PF}(G) : G \in \mathcal{G}(n,m)\}$.

Proposition 7 tells us that, if we want to compare aggregate payoffs of any two networks G and G', we can compare their largest real eigenvalues, $\lambda_{\mathrm{PF}}(G)$ and $\lambda_{\mathrm{PF}}(G')$, in the case of strong complementarities λ. Moreover, from Proposition 6 we know that aggregate payoff is highest in the complete network K_n. K_n also has the highest possible largest real eigenvalue, namely $\lambda_{\mathrm{PF}}(K_n) = n-1$ [Cvetkovic and Rowlinson, 1990]. We assume that λ is close to $1/(n-1)$. The closer is the largest real eigenvalue $\lambda_{\mathrm{PF}}(G)$ of a network G to the one of the complete network ($\lambda_{\mathrm{PF}}(K_n) = n-1$), the closer it comes to being efficient. Following these observations we show the ratio of the largest real eigenvalue of stochastically stable networks to $n-1$ for different values of α. We find that for values of α below $1/2$, stochastically stable networks are highly inefficient and a sharp transition occurs for increasing values of α above $1/2$. It is also seen that the transition becomes sharper the larger the network is. This implies that a highly volatile environment and the strong competition of the agents for becoming a hub induces highly inefficient network structures.[41]

7. How Realistic are Nested Split Graphs?

In this section, we would like to discuss two main characteristics of nested-split networks and see how "realistic" they are. First, are real-world networks nested? At the macroscopic level, some real world networks exhibit a high degree of clustering while, coincidentally, their degree distributions show power-law tails. Taken together, these two characteristics indicate a hierarchical organization in the network [Ravasz and Barabási, 2003]. In social and economic (see e.g. May et al. [2008]), as well as biological systems [Bastolla et al., 2009], it has been found that the hierarchical organization of networks can further be characterized by *nestedness*: the neighborhood of a node is contained in the neighborhood of the nodes with higher degrees. In these examples, the extent of nestedness (defined as the fraction of links belonging to the nested structure) was shown to be above 93 % [Saavedra et al., 2008]. A recent study Leskovec et al. [2009] also finds nested core-periphery structures in over 100 large sparse real-world social and information networks.

[40] Stochastically stable networks might not always be connected. Instead, they can have a single connected component and isolated agents. If there are k isolated agents then their contribution to the aggregate payoff is k (with an equilibrium effort equal to one for each agent). We neglect the contribution of these isolated agents because it is negligible for large complementarities, and consider only the connected component of the network.

[41] It can be shown that the largest real eigenvalue can be increased by concentrating all the links in a densely connected core (clique) for fixed values of the number of links m and nodes n [Cvetkovic and Rowlinson, 1990].

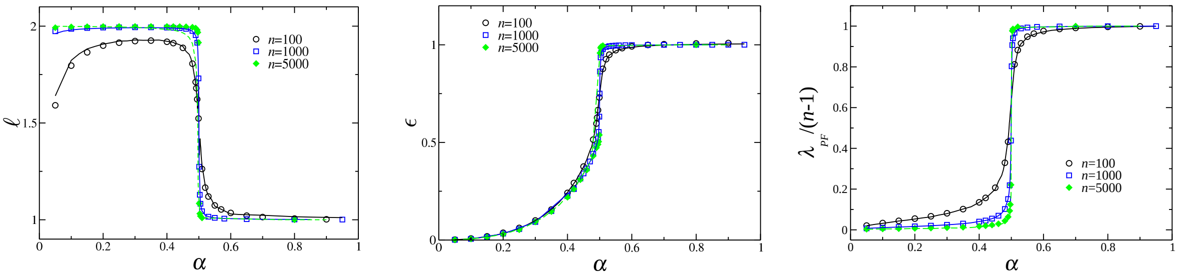


Figure 8: (Left) The characteristic path length ℓ of stationary networks and (middle) the results for the network efficiency ϵ, obtained by recourse of numerical simulations (symbols) and respecting theoretical predictions (lines) of the model. (Right) We show the largest eigenvalue of the adjacency matrix of the stochastically stable network relative to the eigenvalue to the complete graph, which is the efficient network, obtained by recourse of numerical simulations (symbols) and respecting theoretical predictions (lines) of the model for different values of α and $n = 200$ agents. For higher values of α the stochastically stable network comes close to the efficient graph G^* which has a largest real eigenvalue of $\lambda_{\mathrm{PF}}(G^*) = n - 1$.

More generally, as described in the introduction, nestedness appears in various social contexts, including the organization of the New York garment industry [Uzzi, 1996] and in the topology of the Fedwire network [May et al., 2008; Soramaki et al., 2007]. Also, Åkerman and Larsson [2010], who study the evolution of the global arms trade network using a unique dataset on all international transfers of major conventional weapons over the period 1950-2007, find that networks are nested and dissortative in the sense that big countries mainly trade arms with small countries but small countries do not trade with each other. Similarly, using aggregate bilateral imports from 1950 to 2000, De Benedictis and Tajoli [2011] analyze the structure of the world trade network over time, detecting and interpreting patterns of trade ties among countries. They also find that world trade tends to be concentrated among a sub-group of countries and a small percentage of the total number of flows accounts for a disproportionally large share of world trade. Figure 3 in their paper shows a clear core-periphery structure similar to our Figure 3. Indeed, in the 1950s, the trade network was relatively close to that of our Figure 3 when α is small while, in the 1990s, it resembles more to the case when α is high.

Second, nested-split graphs have a diameter equals to 2. Most real-world networks have, in fact, very low diameter. For example, if we consider the Fedwire bank network discussed above [Soramaki et al., 2007], then the average path length (which is the average distance from a node to any other node) is 2.6 while the diameter is around 6 (see their Table 3 on page 324). This means that this interbank payment network exhibits the small world phenomenom common to many networks. Similarly, Åkerman and Larsson [2010] who study arms-trade networks find also small diameters: between 4 and 6. Finally, Banerjee et al. [2010] have detailed network data on 75 villages in rural Karnataka (2-3 hours outside Bangalore in India) before a particular microfinance institution starts working in the villages. They study how the idea of joining microfinance spread through the network. In this dataset, the social network consists of friends and relatives and from whom they borrow or lend small amount of money. The average path length is 1.97 and the diameter is 3. All these networks have small diameters but, of course, not equal exactly to 2. We can extend our model to obtain networks with diameters closer to 6 than to 2 [see König et al., 2010], but adding these features would reduce the mathematical elegance of our formulation in terms of nested-split graphs. At the end of this section, we provide some simulations when $\zeta > 0$ and show that the networks are still nested-split graphs but the diameter is bigger than 2.

To sum up, nested-split graphs have a much more regular structure than the complex networks we observe in the real world but are easy to study, they are the result of endogenous rational actions and, even if different, they have most of the properties of real-world networks. In this respect, nested-split graphs are a good approximation of real-world networks. Also, because nested-split graphs have a simple structure, our model can be easily extended to study different issues. For example, König and Tessone [2011] show that our model can be applied not only to an economic context, but also to a variety of models studied in the physics literature, ranging from the analysis of ecological systems to physical synchronization processes being coupled to network dynamics. They extend our model by introducing heterogeneous selection probabilities of the nodes depending on the number of links they already have, derive the dynamics of the degree distribution in the continuous limit and analyze its properties. They show that the stationary degree distribution is given by a double-power law with a flexible exponent.[42]

We have seen that the stochastically stable network, i.e. the network in the support of the stationary distribution obtained when ζ goes to zero, is a nested-split graph. In contrast, when ζ tends to infinity, links will be created and deleted at random and therefore we will obtain a Poisson random graph.[43] For intermediary values of ζ, the Markov chain will still be ergodic but the stationary distribution will be difficult to characterize. This clear transition can be observed in panels (a) to (c) in Figure 9 (right). We have run numerical simulations of our model by characterizing the global network properties of the stationary networks for $\zeta > 0$. In that case, we see that the stationary states display characteristics more commonly found in real-world networks, like the average path length, which takes small values, but larger than $\ell = 2$ (See Figure 9 (left)). This means that when ζ is not equal to zero but close to it, we still have a nested-split graph structure with a path length larger than 2 but still relatively small. Indeed, in the simulations displayed in the middle of Figure 9, when $\zeta = 10$, the average path length ℓ is equal to 3.

8. Estimating the Model's Parameters

In this section we would like to provide real-world evidence of our model and estimate the model's parameters for four different empirical data sets.

The first network we analyze is the network of Austrian banks in the year 2008 [cf. Boss et al., 2004]. Links in the network represent exposures between Austrian-domiciled banks on a non-consolidated basis (i.e. no exposures to foreign subsidiaries are included). We obtain a sample of $n = 770$ banks with $m = 2454$ links between them and an average degree of $\bar{d} = 20.54$. The degree variance is $\sigma_d^2 = 1273.22$. The largest connected component comprises 768 banks, which is 99.7% of the total of banks, and it is illustrated in Figure 10. The network is highly clustered with an average clustering coefficient of $C = 0.75$ and shows a low average path length of $\ell = 2.27$ in the connected component. The network of banks is dissortative with a monotonic decreasing nearest neighbor connectivity d_{nn} with

[42] Such distributions have been observed for example in online social networks [Gjoka et al., 2010].
[43] See Equation (6).

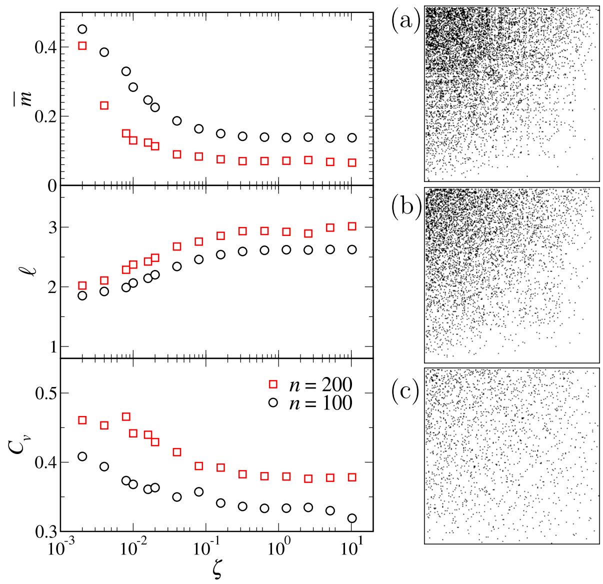


Figure 9: Results of numerical simulations of the model introduced in Definition 2, where the rates of link creation (decay) are Gumbel distributed on the increase (resp. decrease) of payoff of the agents adjacent to the link. In the left panels (from top to bottom) the network density $\bar{m}$, the average path-length (ℓ) and the degree of centralization for the eigenvector centrality C_v are plotted as a function of ζ, the parameter of the Gumbel distribution. The three plots on the right column depict snapshots of the adjacency matrix once the stationary regime has been reached (all the global measures do not change over time, but fluctuate around the stationary state). In these snapshots, the parameters are (a) $\zeta = 0.002$, (b) $\zeta = 0.004$, and (c) $\zeta = 0.08$, respectively.

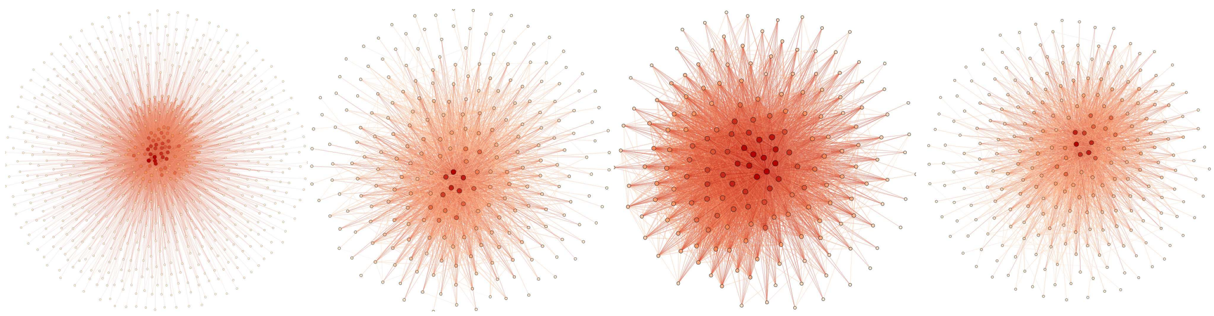


Figure 10: Adjacency matrices for the Austrian banking network, the global network of banks obtained from the Bank of International Settlements (BIS) locational statistics, the GDP trade network and the arms trade network (from left to right). The shade and size of the nodes indicate their eigenvector centrality. The GDP trade network is much more dense than the network of banks and the network of arms trade. All four networks show a core of densely connected nodes.

degree (see Figure 12) and an assortativity coefficient of $\gamma = -0.51$ [Newman, 2002].[44]

Secondly, we consider the global banking network in the year 2011 obtained from the Bank of International Settlements (BIS) locational statistics on exchange-rate adjusted changes in cross-border bank claims [cf. Minoiu and Reyes, 2011]. BIS locational statistics are compiled on the basis of residence of BIS reporting banks and cover the cross-border positions of all banks domiciled in the reporting area, including positions with respect to foreign affiliates, loans, deposits, debt securities, and other assets provided by banks. We obtain a network with $n = 239$ nodes and $m = 2454$ links between them. An illustration can be seen in Figure 10. The average degree of the network is $\bar{d} = 20.54$ and we observe a high degree variance of $\sigma_d^2 = 1273.22$. Almost all nodes are contained in the largest connected component (93%). The average clustering coefficient is high and measures $C = 0.81$. The network is characterized by a short average path length (in the connected component) of $\ell = 1.95$. The network is dissortative with a coefficient of $\gamma = -0.76$. and a monotonic decreasing average nearest neighbor connectivity d_{nn} (see Figure 12).

The third empirical network we consider is the network of trade relationships between countries in the year 2000 [cf. Gleditsch, 2002]. The trade network is defined as the network of import-export relationships between countries in a given year in millions of current-year U.S. dollars. We refer to this network as the GDP network. We construct an undirected network in which a link is present between two countries if either one has exported to the other country. The trade network contains $n = 196$ nodes, $m = 4138$ links, has an average degree of $\bar{d} = 42.22$ and a degree variance of $\sigma_d^2 = 1524.16$. The network consists of a giant component with 181 nodes, encompassing 92% of all nodes in the network. The network is shown in Figure 10. The network of trade is highly clustered with $C = 0.73$, a short average path length of $\ell = 2.25$, dissortative $\gamma = -0.40$ and has a decreasing average nearest neighbor connectivity d_{nn} (see Figure 12).

Fourth, we consider the network of arms trade between countries [cf. Åkerman and Larsson, 2010]. We use data obtained from the SIPRI Arms Transfers Database holding information on all international transfers between countries of seven categories of major conventional weapons accumulated from 1950 to 2010. A link in the network represents a recipient or supply relationship of arms between two countries during this period. We obtain a network with $n = 246$ nodes and $m = 2245$ links. The average degree is $\bar{d} = 18.25$ and the degree variance is $\sigma_d^2 = 589.97$. An illustration can be seen in Figure 10. The network is connected and has an average path length of $\ell = 2.25$. The average clustering coefficient is $C = 0.48$ and the average nearest neighbor connectivity d_{nn} is decreasing with degree (see Figure 12), with an assortativity coefficient of $\gamma = -0.39$.

Observe that our theoretical model is general enough to encompass all these networks. For example, consider our payoff function of Equation (1) defined in Section 3.1. Then, for the trade network, it suffices to interpret x_i as the volume of trade for country i and have $a_{ij} = 1$ between countries i and j when either one has exported to the other country. Similarly, for the bank network, x_i would be the volume of loans each bank i provides to other banks in the network and $a_{ij} = 1$ when bank i provides a loan to bank j. In that case, we need to consider directed networks instead of undirected networks but this does not affect any of our results.

All these four real-world networks are of similar size, show short average path lengths

[44] See also Footnote 37 for the definition of γ.

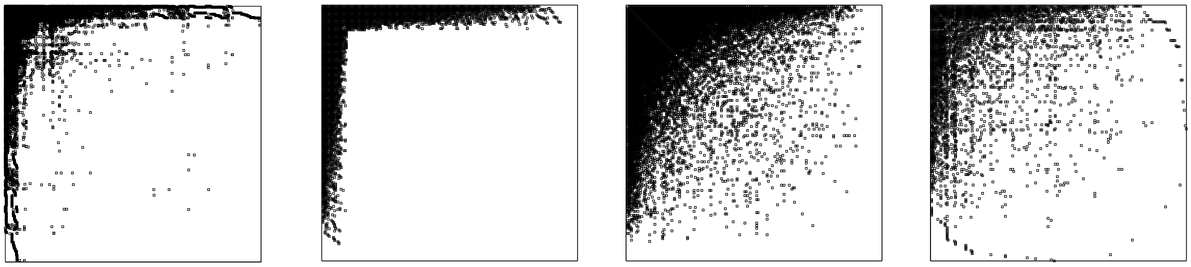


Figure 11: Adjacency matrices (sorted by the eigenvector centralities of the nodes) for the Austrian banking network, the global network of banks obtained from the Bank of International Settlements (BIS) locational statistics, the GDP trade network and the arms trade network (from left to right). All adjacency matrices are significantly nested.

of around 2, are dissortative and have a monotonic decreasing average nearest neighbor connectivity. They also show a relatively high clustering and the clustering degree distribution is decreasing with the degree (see Figure 12). An important feature of these networks is that they all show a high degree of *nestedness*. This can be witnessed from the adjacency matrices depicted in Figure 11, which resemble the nested matrices we derive from our theoretical model (see Figure 2). Similarly, when we compare the networks simulated from our model (see Figure 3) and the ones described in real-world networks (Figure 10), they are relatively similar (in terms of a clear core-periphery structure indicating nestedness).[45] A more quantitative measure of nestedness confirms this observation by showing that all networks are significantly nested. We have computed the degree of nestedness by calculating the *matrix temperature* T_n using the `BINMATNEST` algorithm proposed by Rodríguez-Gironés and Santamaría [2006]. Typically, the lower the temperature T_n, the higher the degree of nestedness. For the network of Austrian banks we obtain $T_n = 0.05$, for the network of banks from BIS statistics we get $T_n = 0.75$, for the GDP trade network we obtain $T_n = 7.26$, and for the arms trade network $T_n = 1.72$. This indicates that the networks of banks have the highest degree of nestedness. However, comparing T_n for matrices of different density and size can be difficult. Instead, one can compute the probability of a certain degree of nestedness being generated at random. For all the networks considered we obtain a p-value not distinguishable from zero (using 500 null matrices) showing that all empirical networks are significantly nested.[46]

We then estimate the main parameters $\mathbf{\Theta} \equiv (\alpha, \zeta)$ of our model by using the Likelihood-Free Markov Chain Monte Carlo (LF-MCMC) algorithm suggested by Marjoram et al.

[45] We have performed a k-core decomposition of the empirical networks. A k-core is a maximal subnetwork in which all nodes have a degree of at least k with the other nodes in the subnetwork. Examining the k-cores with increasing values of k does not split the network into separate components. This is another indicator for the nested structure observed in these networks.

[46] See Rodríguez-Gironés and Santamaría [2006] for further details of the `BINMATNEST` algorithm.

[2003].[47] The purpose of this algorithm is to estimate the parameter vector $\mathbf{\Theta}$ of our model on the basis of the summary statistics $\mathbf{S} \equiv (\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3)_{n\times 3}$, where $\mathbf{S}_1 \equiv (P(d))_{d=0}^{n-1}$, $\mathbf{S}_2 \equiv (C(d))_{d=0}^{n-1}$, $\mathbf{S}_3 \equiv (d_{nn}(d))_{d=0}^{n-1}$ are the degree distribution, the clustering degree distribution and the average nearest neighbor degree distribution, respectively. Moment conditions are obtained from the Euclidean distances $\Delta(\mathbf{S}_i, \mathbf{S}_i^o) \equiv \sqrt{\sum_{j=1}^{n}\left(S_{i,j} - S_{i,j}^o\right)^2}$ for each statistic $\mathbf{S}_i$ (generated by the algorithm) and its observed value $\mathbf{S}_i^o$. The algorithm generates a Markov chain which is a sequence of parameters $\{\mathbf{\Theta}_s\}_{s=1}^S$ with a stationary distribution that approximates the distribution of the parameter values $\mathbf{\Theta}$ conditional on the observed statistic $\mathcal{S}^o$.[48] Since this estimation algorithm would require the computation of the Bonacich centrality an extensive number of times, we assume that the complementarity parameter λ is small such that we can approximate the Bonacich centrality by the degree centrality when simulating the network formation process.[49] Note also that the reported estimates of ζ hold only up to a scaling factor, which depends on the choice of λ. Hence, only the relative values of ζ between different samples is meaningful but not its absolute value.

The estimated parameter values are shown in Table 1. We observe that the estimates for ζ are higher for the network of GDP trading countries and the network of arms trade than the corresponding estimates for the networks of banks. This confirms our intuition from Figure 9 (right), where we have seen that with increasing values of ζ stationary networks become less nested (and we obtain a random graph as $\zeta \to \infty$), and the lower values for the matrix temperature T_n for these networks (see also the adjacency matrices in Figure 11). Hence, our estimates support our earlier observation that the networks of banks have a higher degree of nestedness than the networks of trade relationships between countries.

Moreover, Figure 12 shows the empirical distributions (squares) and typical simulated ones (circles) for the bank network, the network of GDP trade and the arms trade network. The comparison of observed and the simulated distributions shown in Figure 12 indicate that the model can relatively well reproduce the observed empirical networks, even though the model is parsimoniously parameterized in relying only on two exogenous variables α and ζ. The fit seems to be best for the networks of banks, which also shows the most distinct nestedness pattern (see Figure 11).

9. Robustness Analysis

In our model, we describe a dynamic process incorporating both the play of a network game and the endogenous formation of the network. A striking finding is that, starting

[47]See Sisson and Fan [2011] for an introduction to LF-MCMC and Chib [2001] for a general discussion of MCMC approaches in econometrics.

[48]For the implementation of the algorithm we have chosen an initial uniform (prior) parameter distribution. The proposal distribution is a normal distribution. During the "burn-in" phase [Chib, 2001], we consider a monotonic decreasing sequence of thresholds with appropriately chosen values from careful numerical experimentation. For the Austrian banking network we have chosen a burn-in period of 1000 steps, while for the network of GDP trade we have used a period of 3000.

[49]The Bonacich centrality is defined by $b_i(G,\lambda) = \sum_{k=0}^{\infty}\lambda^k\left(\mathbf{A}^k\cdot\mathbf{u}\right)_i = 1 + \lambda d_i + \lambda^2\sum_{j\in\mathcal{N}_i} d_j + \lambda^3\sum_{j\in\mathcal{N}_i}\sum_{k\in\mathcal{N}_j} d_k + \ldots = 1 + \lambda d_i + \lambda^2\sum_{j\in\mathcal{N}_i} d_j + O(\lambda^3)$. Marginal payoff from forming a link ij for agent i can then be written as $\pi_i^*(G\oplus ij,\lambda) - \pi_i^*(G,\lambda) = \frac{\lambda(2+\lambda)}{2} + \frac{\lambda^2}{2}d_i(d_i+1) + \lambda^2 d_j + O(\lambda^3)$. When computing marginal payoffs from forming a link (and the decay rates) we ignore terms of the order $O(\lambda^3)$.

Table 1: Estimation of the model parameters $\theta \in \mathbf{\Theta} = (\alpha, \zeta)$ for the Austrian network of banks, the global network of banks obtained from the Bank of International Settlements (BIS) locational statistics, the network of GDP trading countries and the arms trade network.[a] The table shows simulated averages of the parameters and their standard deviations,[b] after the chain has converged.[c]

	Austrian Bank Network				Global Bank Network				GDP Trade Network				Arms Trade Network			
θ	μ_θ	σ_θ	τ_θ	$p_\theta(S)$	μ_θ	σ_θ	τ_θ	$p_\theta(S)$	μ_θ	σ_θ	τ_θ	$p_\theta(S)$	μ_θ	σ_θ	τ_θ	$p_\theta(S)$
α	0.45	0.00	29.63	0.99	0.45	0.01	10.65	0.91	0.44	0.00	6.49	0.96	0.44	0.00	50.67	0.99
ζ	1.13	0.06	68.98	0.99	1.70	0.18	91.64	0.96	20.51	0.32	143.30	0.89	24.87	0.16	657.82	0.93
n	770				239				196				246			
S	3000				700				2000				5000			

[a] τ_θ is the integrated autocorrelation time, which should be much smaller than the number S of iterations of the Markov chain algorithm used to compute the parameter estimates [Sokal, 1996].

[b] μ_θ is the mean and σ_θ is the simulation standard deviation calculated from batch means (of length 10) for each parameter $\theta \in \mathbf{\Theta}$ [Chib, 2001].

[c] $p_\theta(S)$ is the p-value associated with Geweke's spectral density diagnostic indicating the convergence of the chain [Brooks and Roberts, 1998; Geweke, 1992]. The number S of iterations of the chain have been chosen for each data set individually such that reasonably high values of $p_\theta(S)$ are obtained.

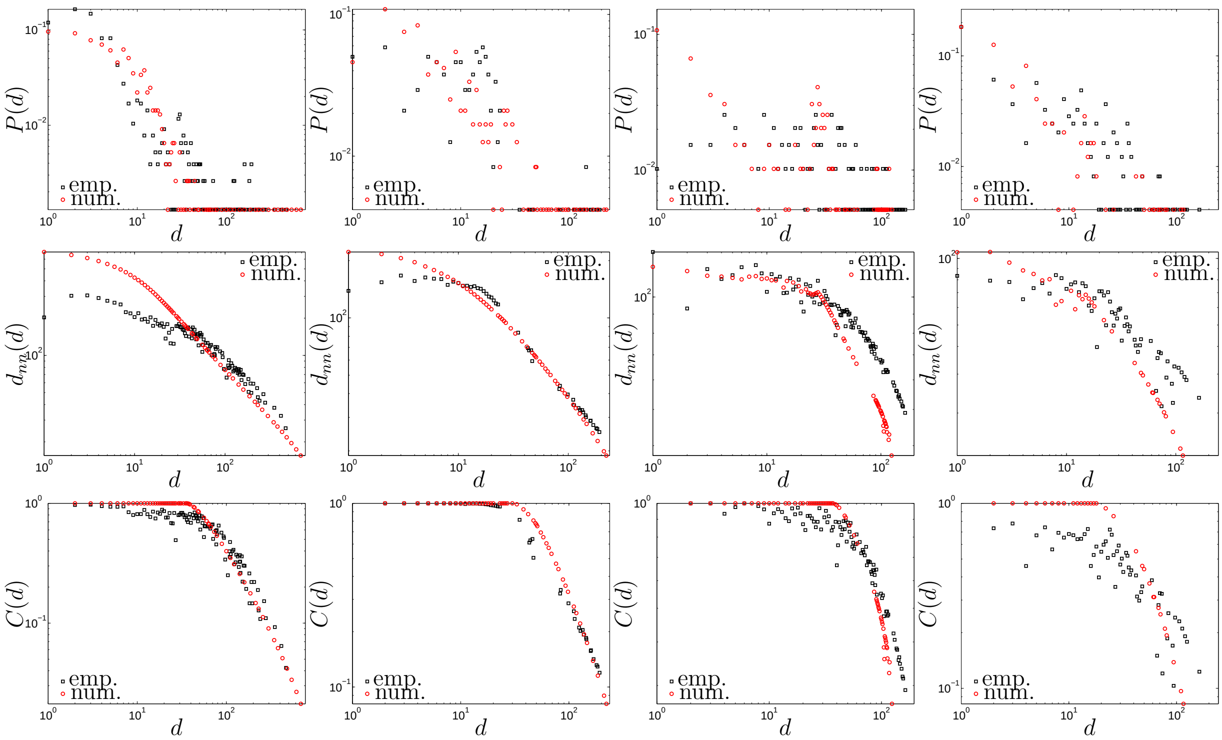


Figure 12: The empirical (□) and an exemplary simulated (○) degree distribution $P(d)$, average nearest neighbor degree $d_{nn}(d)$ and clustering degree distribution $C(d)$ for the Austrian banking network (first column), the network of banks obtained from the Bank of International Settlements (BIS) (second column), the GDP trade network (third column) and the arms trade network (fourth column).

from any arbitrary graph,[50] the process converges to a nested split graph in the limit of vanishing noise. We like to show that this characterization is *not* an artifact of a very specific protocol of network formation but is quite general.

First, in the model presented in this paper, using Ballester et al. [2006], we give a micro-foundation of why agents choose to create a link with the agent who has the highest Bonacich centrality in the network. By doing so we impose a specific structure of the utility function (i.e. a linear quadratic structure; see Equation (1)) and a condition on the largest real eigenvalue of the graph (i.e. $\lambda < 1/\lambda_{\mathrm{PF}}$; see Theorem 1). In fact, all our results, in particular the fact that the network converges to a nested split graph, hold if we take a general utility function (i.e. no specific structure), that is increasing in their Bonacich centrality. Moreover, imagine that agents do not choose effort $\mathbf{x}$ but just create links following the network formation mechanism described in Definition 2, where their utility is given by the sum of the current Bonacich centrality of their neighbors, then all our results still hold. As a consequence of Corollary 11 (see Appendix B.2.5), this is also true for any other centrality measure we have considered. Moreover, we could also define as utility "information centrality" introduced by Stephenson and Zelen [1989], where the payoff of agents is given by the information they receive along different paths in the network. If we let agents choose the importance (or weights, see Stephenson and Zelen [1989]) of the paths such that they receive maximum information (similar to choosing efforts $\mathbf{x}$) then all our results hold without making any assumption on the eigenvalue of the network.[51]

Second, we consider in C the same model as in Section 3 but we assume that the agent who wants to form a link must pay a cost $c \geq 0$. Proposition 8 in C shows that if c is less than $\lambda(2-\lambda)/(2(1-\lambda)^2)$ then agents payoffs are increasing from forming a link. As a result, even when requiring that links are only formed if they profitable, the emerging network will always be a nested split graph. In other words, even with costly link formation, all our results hold as long as c is not too large.

Third, in Section 5, for the nodes in the dominating subsets, we obtain a power-law degree distribution with an exponent of minus one. We can extend our model to obtain a degree distribution with an arbitrary power-law tail by making the probability of selecting an agent depending on the number of links she already has, while preserving the nested structure of the network she is embedded in. This extension is studied in König and Tessone [2011], where also a number of applications of our model to various physical and ecological systems is discussed.

Fourth, with our network formation game, we always obtain negative degree-degree correlations (i.e. our networks are dissortative). In König et al. [2010], we extend our game by including capacity constraints in the number of links an agent can maintain and a random search mechanism for new linking partners when an agent refuses to accept a new link due to her capacity constraints. We find that by introducing capacity constraints and random search, stationary networks can become assortative. Thus, we are able to reproduce all topological properties of empirically observed social and economic networks. The emergence of assortativity and positive degree-correlations, respectively, can be explained by considering limitations in the number of links an agent can maintain. This may be of

[50] See Proposition 3 in Section 4.

[51] The same reasoning as in the proof of Proposition 1 that applies for walks also applies for paths. This fact can be used to show that the agent with the highest degree in a nested split graph has the highest information centrality.

particular relevance for social networks and give an explanation for the distinction between assortative social networks and dissortative technological networks suggested by Newman [2002].

Finally, we have assumed myopic agents, i.e. agents maximize their current utility. Assume, instead, that agents create and delete links based on their discounted life time expected utility. They will still create links with the agent who has the highest Bonacich centrality in the network since this agent has a higher probability than any other agent in the network to have the highest Bonacich centrality in the future (at any period of time). In other words, there exists a farsighted equilibrium which has the same properties as ours, i.e., the network at any period of time is a nested-split graph.

10. Conclusion

In this paper we propose a dynamic network formation model to explain the observed nestedness in various real-world networks. Agents exert effort in some activities and create links with other agents while the network is exposed to a volatile environment. We find that the stochastically stable networks are nested split graphs. We completely determine the topological properties of the stochastically stable networks and show that they match features exhibited by real-world networks. We estimate the model using data from four different real-world networks and show that it fits well the observed networks.

The model and analysis provide an explanation of the importance of nestedness in real-world networks. Because agents tend to link to highly connected agents, a core-periphery network structure tends to emerge in equilibrium as a result of this network formation process. This is particularly true for networks of banks since small banks interact with large banks for security, lower liquidity risk and lower servicing costs, and large banks interact with small banks in part because they can extract a higher premium for services and can accommodate more risk. This is certainly less true for co-authorship networks, which tend to be non-nested and assortative [cf. Goyal et al., 2006]. There is certainly much more work to be done in this area but we believe that our model, by providing a microfoundation to the link-formation process and unifying the strategic-based and the random-based network formation models, provides a first step towards a better understanding of the structure of real-world networks.

References

Åkerman, A., Larsson, A., 2010. The global arms trade network 1950–2007. Stockholm University Working Paper No. 2010:2.

Albert, R., Barabási, A.-L., Jan 2002. Statistical mechanics of complex networks. Rev. Mod. Phys. 74 (1), 47–97.

Anderson, S. A., De Palma, A., Thisse, J., 1992. Discrete Choice Theory of Product Differentiation. MIT Press.

Aouchiche, M., Bell, F., Cvetkovic, D., Hansen, P., Rowlinson, P., Simic, S., Stevanovic, D., 2008. Variable neighborhood search for extremal graphs, 16. some conjectures related to the largest eigenvalue of a graph. European Journal of Operational Research 191 (3), 661–676.

Arenas, À., Cabrales, A., Danon, L., Díaz-Guilera, A., Guimerà, R., Vega-Redondo, F., 2010. Optimal information transmission in organizations: Search and congestion. Review of Economic Design 14 (1), 75–93.

Baker, G., Gibbons, R., Murphy, K., 2008. Strategic alliances: Bridges between islands of conscious power? Journal of the Japanese and International Economies 22 (2), 146–163.

Bala, V., Goyal, S., 2000. A noncooperative model of network formation. Econometrica 68 (5), 1181–1230.

Ballester, C., Calvó-Armengol, A., Zenou, Y., 2006. Who's who in networks. wanted: The key player. Econometrica 74 (5), 1403–1417.

Banerjee, A., Chandrasekhar, A., Duflo, E., Jackson, M., 2011. The diffusion of microfinance. Unpublished manuscript, MIT.

Barabasi, A.-L., Albert, R., 1999. Emergence of scaling in random networks. Science 286, 509–512.

Bastolla, U., Fortuna, M., Pascual-García, A., Ferrera, A., Luque, B., Bascompte, J., 2009. The architecture of mutualistic networks minimizes competition and increases biodiversity. Nature 458 (7241), 1018–1020.

Bavelas, A., 1948. A mathematical model for group structures. Human Organization 7, 16–30.

Beauchamp, M. A., 1965. An improved index of centrality. Behavioral Science 10, 161–163.

Blume, L., 2003. How noise matters. Games and Economic Behavior 44 (2), 251–271.

Boje, D. M., Whetten, D. A., 1981. Effects of organizational strategies and contextual constraints on centrality and attributions of influence in interorganizational networks. Administrative Science Quarterly 26 (March), 378–395.

Bolland, J., 1988. Sorting out centrality: An analysis of the performance of four centrality models in real and simulated networks. Social Networks 10 (3), 233–253.

Bonacich, P., 1987. Power and centrality: A family of measures. American Journal of Sociology 92 (5), 1170.

Bonacich, P., Lloyd, P., 2001. Eigenvector-like measures of centrality for asymmetric relations. Social Networks 23 (3), 191–201.

Borgatti, S. P., 2005. Centrality and network flow. Social Networks 27 (2), 55–71.

Boss, M., Elsinger, H., Summer, M., Thurner, S., 2004. Network topology of the interbank market. Quantitative Finance 4 (6), 677–684.

Bramoullé, Y., Kranton, R., July 2007. Public goods in networks. Journal of Economic Theory 135 (1), 478–494.

Bramoullé, Y., Kranton, R., D'Amours, M., 2011. Strategic interaction and networks. Unpublished manuscript, Duke University.

Bramoullé, Y., Lopez, D., Goyal, S., Vega-Redondo, F., 2004. Social interaction in anti-coordination games. International Journal of Game Theory 33 (1), 1–19.

Brass, D. J., 1984. Being in the right place: A structural analysis of individual influence in an organization. Administrative Science Quarterly 29 (4), 518–539.

Brooks, S., Roberts, G., 1998. Assessing convergence of markov chain monte carlo algorithms. Statistics and Computing 8 (4), 319–335.

Brualdi, R. A., Hoffman, A. J., 1985. On the spectral radius of 0-1 matrices. Linear Algebra and Application 65, 133–146.

Cabrales, A., Calvó-Armengol, A., Zenou, Y., 2011. Social interactions and spillovers. Games and Economic Behavior 72, 339–360

Calvó-Armengol, A., de Martí, J., 2009. Information gathering in organizations: equilibrium, welfare, and optimal network structure. Journal of the European Economic Association 7 (1), 116–161.

Calvó-Armengol, A., Patacchini, E., Zenou, Y., 2009. Peer effects and social networks in education. Review of Economic Studies 76, 1239–1267.

Calvó-Armengol, A., Zenou, Y., 2004. Social Networks and Crime Decisions: The Role of Social Structure in Facilitating Delinquent Behavior. International Economic Review 45 (3), 939–958.

Chib, S., 2001. Markov chain monte carlo methods: computation and inference. Handbook of econometrics 5, 3569–3649.

Cohen-Cole, E., Patacchini, E., Zenou, Y. 2011. Systemic risk and network formation in the interbank market. CEPR Discussion Paper No. 8332.

Cook, K. S., Emerson, R. M., Gillmore, M. R., Yamagishi, T., 1983. The distribution of power in exchange networks: Theory and experimental results. The American Journal of Sociology 89 (2), 275–305.

Corbo, J., Calvó-Armengol, A., Parkes, D., 2006. A study of nash equilibrium in contribution games for peer-to-peer networks. SIGOPS Operation Systems Review 40 (3), 61–66.

Cvetkovic, D., Rowlinson, P., 1990. The largest eigenvalue of a graph: A survey. Linear and Multinilear Algebra 28 (1), 3–33.

Cvetkovic, D., Rowlinson, P., Simic, S., 1997. Eigenspaces of Graphs. Vol. 66. Cambridge University Press, Encyclopedia of Mathematics.

De Benedictis, L., Tajoli, L., 2011. The world trade network. The World Economy 34 (8), 1417–1454.

De Masi, G., Gallegati, M., 2007. Econophysics of Markets and Business Networks. Springer Milan, Ch. Debt-credit Economic Networks of Banks and Firms: the Italian Case, pp. 159–171.

Debreu, G., Herstein, I. N., 1953. Nonnegative square matrices. Econometrica 21 (4), 597–607.

Dodds, P. S., Watts, D. J., Sabel, C. F., 2003. Information exchange and robustness of organizational networks. Proceedings National Academy of Sciences 100 (21), 12516–12521.

Dupouet, O., Yildizoglu, M., 2006. Organizational performance in hierarchies and communities of practice. Journal of Economic Behavior and Organization 61, 668–690.

Dutta, B., Ghosal, S., Ray, D., 2005. Farsighted network formation. Journal of Economic Theory 122 (2), 143–164.

Ehrhardt, G., Marsili, M., Vega-Redondo, F., 2006. Diffusion and growth in an evolving network. International Journal of Game Theory 34 (3).

Ehrhardt, G., Marsili, M., Vega-Redondo, F., 2008. Networks emerging in a volatile world, economics Working Papers.

Ellison, G., 2000. Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution. Review of Economic Studies 67 (1), 17–45.

Ethier, S., Kurtz, T., 1986. Markov processes. Wiley.
Feri, F., 2007. Stochastic stability in networks with decay. Journal of Economic Theory 135 (1), 442–457.
Foster, D., Young, P., 1990. Stochastic evolutionary game dynamics. Theoretical Population Biology 38 (2), 219–232.
Freeman, L., 1979. Centrality in social networks: Conceptual clarification. Social Networks 1 (3), 215–239.
Freeman, L. C., 1977. A set of measures of centrality based on betweenness. Sociometry 40, 35–41.
Galeotti, A., Goyal, S., 2010. The law of the few. The American Economic Review 100 (4), 1468–1492.
Galeotti, A., Goyal, S., Jackson, M., Vega-Redondo, F., Yariv, L., 2010. Network games. Review of Economic Studies 77 (1), 218–244.
Gallager, R., 1996. Discrete stochastic processes. Springer.
Geweke, J., 1992. Bayesian Statistics 4. Oxford University Press, Ch. Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments, pp. 169–193.
Gjoka, M., Kurant, M., Butts, C. T., Markopoulou, A., March 2010. Walking in Facebook: A Case Study of Unbiased Sampling of OSNs. In: Proceedings of IEEE INFOCOM '10. San Diego, CA.
Gleditsch, K., 2002. Expanded trade and GDP data. Journal of Conflict Resolution 46 (5), 712.
Goyal, S., Joshi, S., 2003. Networks of collaboration in oligopoly. Games and Economic Behavior 43 (1), 57–85.
Goyal, S., van der Leij, M., Moraga-Gonzalez, J. L., 2006. Economics: An emerging small world. Journal of Political Economy 114 (2), 403–412.
Goyal, S., Vega-Redondo, F., 2005. Network formation and social coordination. Games and Economic Behavior 50 (2), 178–207.
Grassi, R., Stefani, S., Torriero, A., 2007. Some new results on the eigenvector centrality. The Journal of Mathematical Sociology 31 (3), 237 – 248.
Grewal, R., Lilien, G. L., Mallapragada, G., 2006. Location, location, location: How network embeddedness affects project success in open source systems. Management Science 52 (7), 1043–1056.
Grimmett, G., Stirzaker, D., 2001. Probability and random processes. Oxford University Press.
Grossman, J., 2002. The evolution of the mathematical research collaboration graph. Congressus Numerantium, 201–212.
Guimerà, R., Díaz-Guilera, A., Giralt, F., Arenas, A., 2006. The real communication network behind the formal chart: Community structure in organizations. Journal of Economic Behavior and Organization 61, 653–667.
Guimerà, R., Díaz-Guilera, A., Vega-Redondo, F., Cabrales, A., Arenas, A., Nov 2002. Optimal network topologies for local search with congestion. Physical Review Letters 89 (24), 248701.
Hagberg, A., Swart, P., Schult, D., 2006. Designing threshold networks with given structural and dynamical properties. Physical Review E 74 (5), 56116.
Hagedoorn, J., May 2002. Inter-firm R&D partnerships: an overview of major trends and patterns since 1960. Research Policy 31 (4), 477–492.
Harrigan, K., 1988. Strategic alliances and partner asymmetries. Management international review 28 (4), 53–72.
Haynie, D., 2001. Delinquent peers revisited: Does network structure matter? American Journal of Sociology 106, 1013–1057.
Hofbauer, J., Sandholm, W., 2007. Evolution in games with randomly disturbed payoffs. Journal of Economic Theory 132 (1), 47–69.
Hu, H., Wang, X., 2009. Disassortative mixing in online social networks. Europhysics Letters 86, 18003.
Huberman, Bernardo, A., Hogg, T., 1995. Communities of practice: Performance and evolution. Computational and Mathematical Organization Theory 1 (1), 73–92.
Jackson, M. O., 2007. The economics of social networks. In: Advances in Economics and Econometrics. Theory and Applications, Ninth World Congress. Vol. 1. pp. 1–56.
Jackson, M. O., 2008. Social and Economic Networks. Princeton University Press.
Jackson, M. O., Rogers, B. W., 2007. Meeting strangers and friends of friends: How random are social networks? American Economic Review 97 (3), 890–915.
Jackson, M. O., Watts, A., 2002a. On the formation of interaction networks in social coordination games. Games and Economic Behavior 41 (2), 265–291.
Jackson, M. O., Watts, A., 2002b. The evolution of social and economic networks. Journal of Economic Theory 106 (2), 265–295.
Jackson, M. O., Wolinsky, A., 1996. A strategic model of social and economic networks. Journal of Economic Theory 71 (1), 44–74.
Kandori, M., Mailath, G., Rob, R., 1993. Learning, mutation, and long run equilibria in games. Econometrica: Journal of the Econometric Society, 29–56.
Kemeny, J., Snell, J., 1960. Finite markov chains. Van Nostrand.
Killworth, P., Bernard, H., 1976. Informant accuracy in social network data. Human Organization 35 (3), 269–286.
Kitsak, M., Riccaboni, M., Havlin, S., Pammolli, F., Stanley, H., 2010. Scale-free models for the structure of business firm networks. Physical Review E 81, 036117.
Kogut, B., Urso, P., Walker, G., 2007. Emergent properties of a new financial market: American venture capital syndication, 1960–2005. Management Science 53 (7), 1181–1198.
König, M. D., Tessone, C. J, Zenou, Y., 2010. From Assortative to Dissortative Networks: The Role of Capacity Constraints. Advances in Complex Systems 13 (4), 483–499.
König, M. D., Tessone, C. J., 2011. Network evolution based on centrality. Phys. Rev. E 84, 056108.

Konishi, H., Ray, D., 2003. Coalition formation as a dynamic process. Journal of Economic Theory 110 (1), 1–41.
Latora, V., Marchiori, M., Oct 2001. Efficient behavior of small-world networks. Physical Review Letters 87 (19), 198701.
Leskovec, J., Lang, K., Dasgupta, A., Mahoney, M., 2009. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. Internet Mathematics 6 (1), 29–123.
Liggett, T., 2010. Continuous time Markov processes: an introduction. Vol. 113. American Mathematical Society.
Mahadev, N., Peled, U., 1995. Threshold Graphs and Related Topics. North Holland.
Marjoram, P., Molitor, J., Plagnol, V., Tavaré, S., 2003. Markov chain Monte Carlo without likelihoods. Proceedings of the National Academy of Sciences of the United States of America 100 (26), 15324.
Mas, D., Vignes, A., Weisbuch, G., 2007. Networks and syndication strategies: Does a venture capitalist need to be in the center?, ERMES-CNRS Universite Pantheon Assas Paris II, Working Paper No. 07-14.
May, R., Levin, S., Sugihara, G., 2008. Ecology for bankers. Nature 451 (21), 893–895.
McFadden, D., 1981. Econometric models of probabilistic choice. MIT Press, Cambridge, MA, Ch. Structural Analysis of Discrete Data with Econometric Applications.
Mehra, A., Kilduff, M., Brass, D. J., 2001. The social networks of high and low self-monitors: Implications for workplace performance. Administrative Science Quarterly 46 (1), 121–146.
Mele, A., 2010. Segregation in Social Networks: Theory, Estimation and Policy. Unpublished mauscript.
Minoiu, C., Reyes, J., 2011. A network analysis of global banking: 1978-2009. IMF Working Papers.
Myerson, R. B., 1991. Game Theory: Analysis of Conflict. Harvard University Press.
Newman, M. E. J., 2010. Networks: An Introduction. Oxford University Press.
Newman, M. E. J., Oct 2002. Assortative mixing in networks. Physical Review Letters 89 (20), 208701.
Newman, M. E. J., Feb 2003. Mixing patterns in networks. Physical Review E 67 (2), 026126.
Newman, M. E. J., 2004. Coauthorship networks and patterns of scientific collaboration. Proceedings of the National Academy of Sciences 101 (90001), 5200–5205.
Park, S., Russo, M., 1996. When competition eclipses cooperation: An event history analysis of joint venture failure. Management science, 875–890.
Pastor-Satorras, R., Vazquez, A., Vespignani, A., 2001. Dynamical and correlation properties of the internet. Physical Review Letters 87.
Perry-Smith, J., Shalley, C., 2003. The social side of creativity: A static and dynamic social network perspective. Academy of Management Review 28 (1), 89–106.
Powell, W., Koput, K. W., Smith-Doerr, L., 1996. Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. Administrative Science Quarterly 41 (1), 116–145.
Ravasz, E., Barabási, A., 2003. Hierarchical organization in complex networks. Physical Review E 67 (2), 026112.
Rodríguez-Gironés, M., Santamaría, L., 2006. A new algorithm to calculate the nestedness temperature of presence–absence matrices. Journal of Biogeography 33 (5), 924–935.
Saavedra, S., Reed-Tsochas, F., Uzzi, B., December 2008. A simple model of bipartite cooperation for ecological and organizational networks. Nature 457 (7228), 463–466.
Sabidussi, G., 1966. The centrality index of a graph. Psychometrika 31 (4), 581–603.
Sandholm, W., 2010. Population games and evolutionary dynamics. MIT Press.
Serrano, M. A., Boguñá, M., Jul 2003. Topology of the world trade web. Phys. Rev. E 68 (1), 015101.
Sisson, S., Fan, Y., 2011. Handbook of Markov Chain Monte Carlo. Chapman & Hall/CRC Handbooks of Modern Statistical Methods, Ch. Likelihood-free markov chain monte carlo.
Snijders, T., 2001. The Statistical Evaluation of Social Network Dynamics. Sociological Methodology 31 (1), 361–395.
Snijders, T., Koskinen, J., Schweinberger, M., 2010. Maximum likelihood estimation for social network dynamics. The Annals of Applied Statistics 4 (2), 567–588.
Sokal, A., 1996. Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms.
Soramaki, K., Bech, M., Arnold, J., Glass, R., Beyeler, W., 2007. The topology of interbank payment flows. Physica A: Statistical Mechanics and its Applications 379 (1), 317–333.
Staudigl, M., 2011. Potential games in volatile environments. Games and Economic Behavior 72 (1), 271–287.
Stephenson, K., Zelen, M., 1989. Rethinking centrality: Methods and examples. Social Networks 11, 1–37.
Stroock, D., 2005. An introduction to Markov processes. Vol. 230. Springer Verlag.
Uzzi, B., 1996. The sources and consequences of embeddedness for the economic performance of organizations: The network effect. American sociological review, 674–698.
Uzzi, B., 1997. Social structure and competition in interfirm networks: The paradox of embeddedness. Administrative Science Quarterly 42 (1), 35–67.
Vega-Redondo, F., 2007. Complex Social Networks. Series: Econometric Society Monographs. Cambridge University Press.
Wasserman, S., Faust, K., 1994. Social Network Analysis: Methods and Applications. Cambridge.
Watts, A., 2001. A dynamic model of network formation. Games and Economic Behavior 34 (2), 331–341.
Watts, D. J., Strogatz, S. H., 1998. Collective dynamics of small-world networks. Nature 393, 440–442.
Young, H., 1998. Individual strategy and social structure: An evolutionary theory of institutions. Princeton University Press.
Zachary, W. W., 1977. An information flow model for conflict and fission in small groups. Journal of Anthropological Research 33 (4), 452–473.

Appendix

A. Network Definitions and Characterizations (For Online Publication)

A *network* (*graph*) G is the pair $(\mathcal{N}, \mathcal{E})$ consisting of a set of agents (*vertices* or *nodes*) $\mathcal{N} = \{1, \ldots, n\}$ and a set of edges $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$ (*links*) between them. A link ij is *incident* with the vertex $v \in N$ in the network G whenever $i = v$ or $j = v$. There exists a link between vertices i and j such that $a_{ij} = 1$ if $ij \in \mathcal{E}$ and $a_{ij} = 0$ if $ij \notin L$. The *neighborhood* of an agent $i \in \mathcal{N}$ is the set $\mathcal{N}_i = \{j \in \mathcal{N} : ij \in \mathcal{E}\}$. The *degree* d_i of an agent $i \in N$ gives the number of links incident to agent i. Clearly, $d_i = |\mathcal{N}_i|$. Let $\mathcal{N}_i^{(2)} = \bigcup_{j \in \mathcal{N}_i} \mathcal{N}_j \backslash (\mathcal{N}_i \cup \{i\})$ denote the second-order neighbors of agent i. Similarly, the k-th order neighborhood of agent i is defined recursively from $\mathcal{N}_i^{(0)} = \{i\}$, $\mathcal{N}_i^{(1)} = \mathcal{N}_i$ and $\mathcal{N}_i^{(k)} = \bigcup_{j \in \mathcal{N}_i^{(k-1)}} \mathcal{N}_j \backslash \left(\bigcup_{l=0}^{k-1} \mathcal{N}_i^{(l)} \right)$. A *walk* in G of length k from i to j is a sequence $p = \langle i_0, i_1, \ldots, i_k \rangle$ of agents such that $i_0 = i$, $i_k = j$, $i_p \neq i_{p+1}$, and i_p and i_{p+1} are directly linked, for all $0 \leq p \leq k-1$. Agents i and j are said to be *indirectly linked* in G if there exists a walk from i to j in G. An agent $i \in \mathcal{N}$ is *isolated* in G if $a_{ij} = 0$ for all j. The network G is said to be *empty* when all its agents are isolated.

A *subgraph*, G', of G is the graph of subsets of the agents, $\mathcal{N}(G') \subseteq \mathcal{N}(G)$, and links, $\mathcal{E}(G') \subseteq \mathcal{E}(G)$. A graph G is *connected*, if there is a path connecting every pair of agents. Otherwise G is disconnected. The *components* of a graph G are the maximally connected subgraphs. A component is said to be *minimally connected* if the removal of any link makes the component disconnected.

A *dominating set* for a graph $G = (\mathcal{N}, \mathcal{E})$ is a subset $\mathcal{S}$ of $\mathcal{N}$ such that every node not in $\mathcal{S}$ is connected to at least one member of S by a link. An *independent set* is a set of nodes in a graph in which no two nodes are adjacent. For example the central node in a star $K_{1,n-1}$ forms a dominating set while the peripheral nodes form an independent set.

In a *complete* graph K_n, every agent is adjacent to every other agent. The graph in which no pair of agents is adjacent is the empty graph $\overline{K}_n$. A *clique* $K_{n'}$, $n' \leq n$, is a complete subgraph of the network G. A graph is *k-regular* if every agent i has the same number of links $d_i = k$ for all $i \in \mathcal{N}$. The complete graph K_n is $(n-1)$-regular. The cycle C_n is 2-regular. In a *bipartite graph* there exists a partition of the agents in two disjoint sets V_1 and V_2 such that each link connects an agent in V_1 to an agent in V_2. V_1 and V_2 are independent sets with cardinalities n_1 and n_2, respectively. In a complete bipartite graph K_{n_1,n_2} each agent in V_1 is connected to each other agent in V_2. The *star* $K_{1,n-1}$ is a complete bipartite graph in which $n_1 = 1$ and $n_2 = n-1$.

The *complement* of a graph G is a graph $\bar{G}$ with the same nodes as G such that any two nodes of $\bar{G}$ are adjacent if and only if they are not adjacent in G. For example the complement of the complete graph K_n is the empty graph $\bar{K}_n$.

Let $\mathbf{A}$ be the symmetric $n \times n$ *adjacency matrix* of the network G. The element $a_{ij} \in \{0,1\}$ indicates if there exists a link between agents i and j such that $a_{ij} = 1$ if $ij \in \mathcal{E}$ and $a_{ij} = 0$ if $ij \notin \mathcal{E}$. The k-th power of the adjacency matrix is related to walks of length k in the graph. In particular, $\left(\mathbf{A}^k\right)_{ij}$ gives the number of walks of length k from agent i to agent j. The *eigenvalues* of the adjacency matrix $\mathbf{A}$ are the numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i$ has a nonzero solution vector $\mathbf{v}_i$, which is an *eigenvector* associated with λ_i for $i = 1, \ldots, n$. Since the adjacency matrix $\mathbf{A}$ of an undirected graph G is real and symmetric, the eigenvalues of $\mathbf{A}$ are real, $\lambda_i \in \mathbb{R}$ for all $i = 1, \ldots, n$. Moreover, if $\mathbf{v}_i$ and $\mathbf{v}_j$ are eigenvectors for different eigenvalues, $\lambda_i \neq \lambda_j$, then $\mathbf{v_i}$ and $\mathbf{v}_j$ are orthogonal, i.e. $\mathbf{v}_i^T\mathbf{v}_j = 0$ if $i \neq j$. In particular, $\mathbb{R}^n$ has an orthonormal basis consisting of eigenvectors of $\mathbf{A}$. Since $\mathbf{A}$ is a real symmetric matrix, there exists an orthogonal matrix $\mathbf{S}$ such that $\mathbf{S}^T\mathbf{S} = \mathbf{S}\mathbf{S}^T = \mathbf{I}$ (that is $\mathbf{S}^T = \mathbf{S}^{-1}$) and $\mathbf{S}^T\mathbf{A}\mathbf{S} = \mathbf{D}$, where $\mathbf{D}$ is the diagonal matrix of eigenvalues of $\mathbf{A}$ and the columns of $\mathbf{S}$ are the corresponding eigenvectors. The *Perron-Frobenius eigenvalue* $\lambda_{\text{PF}}(G)$ is the *largest real eigenvalue* of $\mathbf{A}$ associated with G, i.e. all eigenvalues λ_i of $\mathbf{A}$ satisfy $|\lambda_i| \leq \lambda_{\text{PF}}(G)$ for $i = 1, \ldots, n$ and there exists an associated nonnegative eigenvector $\mathbf{v}_{\text{PF}} \geq 0$ such that $\mathbf{A}\mathbf{v}_{\text{PF}} = \lambda_{\text{PF}}(G)\mathbf{v}_{\text{PF}}$. For a connected graph G the adjacency matrix $\mathbf{A}$ has a unique largest real eigenvalue $\lambda_{\text{PF}}(G)$ and a positive associated eigenvector $\mathbf{v}_{\text{PF}} > 0$. There exists a relation between the number of walks in a graph and its eigenvalues. The number of closed walks of length k from a agent i in G to herself is given by $\left(\mathbf{A}^k\right)_{ii}$ and the total number of closed walks of length k in G is $\text{tr}\left(\mathbf{A}^k\right) = \sum_{i=1}^n \left(\mathbf{A}^k\right)_{ii} = \sum_{i=1}^n \lambda_i^k$. We further have that $\text{tr}\left(\mathbf{A}\right) = 0$, $\text{tr}\left(\mathbf{A}^2\right)$ gives twice the number of links in G and $\text{tr}\left(\mathbf{A}^3\right)$ gives six times the number of triangles in G.

B. Topological Properties of Nested Split Graphs (For Online Publication)

In this appendix we discuss in more detail the topological properties of nested split graphs that arise from our network formation process. We first derive several network statistics for nested split graphs. We compute the degree distribution, the clustering coefficient, average nearest neighbor neighbor connectivity and the characteristic path length in a nested split graph. In particular, we show that connected nested split graphs have small characteristic path length, which is at most two. We then analyze different measures of centrality in a nested split graph.[52] From the expressions of these centrality measures we then can show that degree, closeness, eigenvector and Bonacich centrality induce the same ordering of nodes in a nested split graph. If the ordering is not strict, then this holds also for betweenness centrality. As we elaborate in more detail in Section 9 this has important implications for the generality of our model. Finally, for all statistics derived in this section we show that they are all completely determined by the degree partition in a nested split graph.

B.1. Network Statistics

In the following sections we will compute the degree connectivity, the clustering coefficient, assortativity and average nearest nearest neighbor connectivity and the characteristic path length in a nested split graph G as a function of the degree partition $\mathcal{D}$ (introduced in Definition 3).

B.1.1. Degree Connectivity

The nested neighborhood structure of a nested split graph allows us to compute the degrees of the nodes according to a recursive equation that is stated in the next corollary.

Corollary 3. *Consider a nested split graph $G = (\mathcal{N}, \mathcal{E})$ and let $\mathcal{D} = (\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_k)$ be the*

[52]See Wasserman and Faust [1994, Chap. 5.2] for an overview of different measures of centrality.

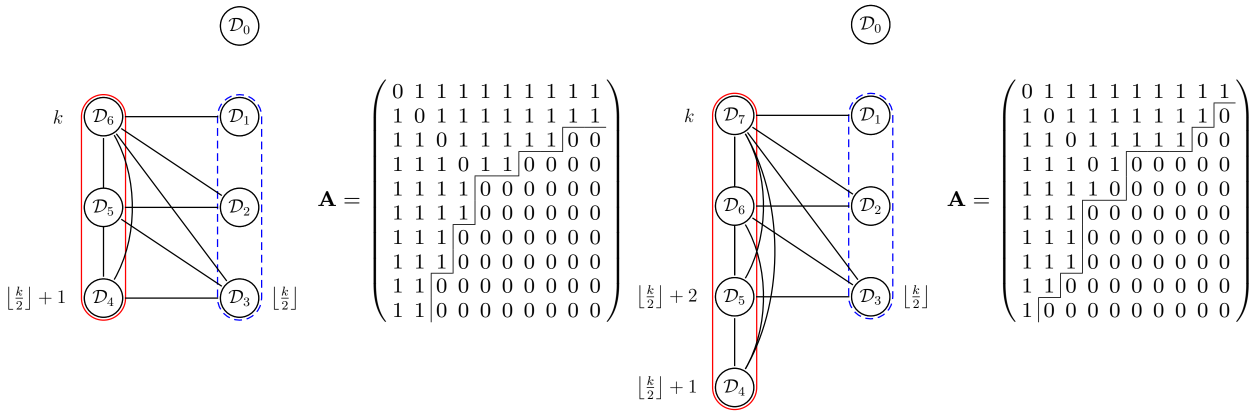


Figure 13: Representation of nested split graphs and their degree partitions $\mathcal{D}$ with corresponding adjacency matrices $\mathbf{A}$. A line between $\mathcal{D}_i$ and $\mathcal{D}_j$ indicates that every node in $\mathcal{D}_i$ is adjacent to every node in $\mathcal{D}_j$. The partitions included in the solid frame ($\mathcal{D}_i$ with $\lfloor \frac{k}{2} \rfloor + 1 \leq i \leq k$) are the dominating subsets while the partitions in the dashed frame ($\mathcal{D}_i$ with $1 \leq i \leq \lfloor \frac{k}{2} \rfloor$) are the independent sets. The figure on the left considers the case of $k = 6$ (even) and the figure on the right the case of $k = 7$ (odd). The illustration follows Mahadev and Peled [1995, p. 11].

degree partition of G. Then $d_u = 0$ if $u \in \mathcal{D}_0$ and for each $u \in \mathcal{D}_i$, $v \in \mathcal{D}_{i-1}$, $i = 1, \ldots, k$, we get

$$d_u = \begin{cases} d_v + |\mathcal{D}_{k-i+1}|, & \text{if } i \neq \lfloor \frac{k}{2} \rfloor + 1, \\ d_v + |\mathcal{D}_{k-i+1}| - 1, & \text{if } i = \lfloor \frac{k}{2} \rfloor + 1, \end{cases} \tag{12}$$

or equivalently

$$d_u = \begin{cases} \sum_{j=1}^{i} |\mathcal{D}_{k+1-j}|, & \text{if } 1 \leq i \leq \lfloor \frac{k}{2} \rfloor, \\ \sum_{j=1}^{i} |\mathcal{D}_{k+1-j}| - 1, & \text{if } \lfloor \frac{k}{2} \rfloor + 1 \leq i \leq k. \end{cases} \tag{13}$$

Equation (12) shows that the neighborhoods of the agents in a nested split graph are nested (see also Definition 4). The degrees of the agents in ascending order of the graph in Figure 13, left, are $2, 3, 4, 5, 7, 9$ while in the graph in Figure 13, right, they are $1, 2, 3, 4, 7, 8, 9$.

B.1.2. Clustering Coefficient

The clustering coefficient $C(u)$ for an agent u is the proportion of links between the agents within her neighborhood $\mathcal{N}_u$ divided by the number of links that could possibly exist between them [Watts and Strogatz, 1998]. It is given by

$$C(u) \equiv \frac{|\{vw : v, w \in \mathcal{N}_u \wedge vw \in \mathcal{E}\}|}{d_u(d_u - 1)/2}. \tag{14}$$

In a nested split graph the clustering coefficient can be derived from the degree partition, as the following corollary shows.

Corollary 4. *Consider a nested split graph $G = (\mathcal{N}, \mathcal{E})$ and let $\mathcal{D} = (\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_k)$ be the degree partition of G. Denote by $S_{\mathcal{D}}^i = \sum_{j=i}^{k} |\mathcal{D}_j|$. Then for each $u \in \mathcal{D}_i$, $i = 0, \ldots, k$, and*

$d_u \geq 2$, the clustering coefficient is given by

$$C(u) = \begin{cases} 0, & \text{if } i = 0, \\ 1, & \text{if } 1 \leq i \leq \left\lfloor \frac{k}{2} \right\rfloor, \\ \frac{1}{d_u(d_u-1)} \left(S_{\mathcal{D}}^{\left\lfloor \frac{k}{2} \right\rfloor+1} - 1 \right) \left[\left(S_{\mathcal{D}}^{\left\lfloor \frac{k}{2} \right\rfloor+1} - 2 \right) + 2|\mathcal{D}_{\left\lfloor \frac{k}{2} \right\rfloor}| \right], & \text{if } i = \left\lfloor \frac{k}{2} \right\rfloor + 1, \quad k \text{ even}, \\ \frac{1}{d_u(d_u-1)} \left(S_{\mathcal{D}}^{\left\lfloor \frac{k}{2} \right\rfloor+1} - 1 \right) \left(S_{\mathcal{D}}^{\left\lfloor \frac{k}{2} \right\rfloor+1} - 2 \right), & \text{if } i = \left\lfloor \frac{k}{2} \right\rfloor + 1, \quad k \text{ odd}, \\ \frac{1}{d_u(d_u-1)} \left[\left(S_{\mathcal{D}}^{\left\lfloor \frac{k}{2} \right\rfloor+1} - 1 \right) \left(S_{\mathcal{D}}^{\left\lfloor \frac{k}{2} \right\rfloor+1} - 2 \right) + \right. & \\ \left. 2 \sum_{j=k-i+1}^{\left\lfloor \frac{k}{2} \right\rfloor} |\mathcal{D}_j| \left(S_{\mathcal{D}}^{k-j+1} - 1 \right) \right], & \text{if } \left\lfloor \frac{k}{2} \right\rfloor + 2 < i \leq k, \end{cases} \tag{15}$$

where d_u is given by Equation (13).

Proof of Corollary 4. Note that for all agents in the independent sets, $u \in \mathcal{D}_i$ with $1 \leq i \leq \left\lfloor \frac{k}{2} \right\rfloor$, the clustering coefficient is one, since their neighbors are all connected among each other. Next, we consider the agents $u \in \mathcal{D}_i$ with $\left\lfloor \frac{k}{2} \right\rfloor + 1 \leq i \leq k$ and degree $d_u = \sum_{j=1}^{i} |\mathcal{D}_{k+1-j}| - 1$. The neighbors of agent u in the dominating subsets are all connected among each other with a total of $\frac{1}{2} \left(\sum_{j=\left\lfloor \frac{k}{2} \right\rfloor+1}^{k} |\mathcal{D}_j| - 1 \right) \left(\sum_{j=\left\lfloor \frac{k}{2} \right\rfloor+1}^{k} |\mathcal{D}_j| - 2 \right)$ links, excluding agent u from the dominating subset. The neighbors of u in the independent sets are not connected. Finally, we consider the links between neighbors for which one neighbor is in a dominating subset and one neighbor is in an independent set. If k is even we get $\sum_{j=k-i+1}^{\left\lfloor \frac{k}{2} \right\rfloor} |\mathcal{D}_j| \left(\sum_{l=k-j+1}^{k} |\mathcal{D}_l| - 1 \right)$ links, excluding agent u in the dominating subset (see Figure 13 (left)). If k is odd there is no such contribution for the agents in the set $\mathcal{D}_{\left\lfloor \frac{k}{2} \right\rfloor+1}$ (see Figure 13 (right)). Putting these contributions together we obtain the clustering coefficient of an agent $u \in \mathcal{D}_i$ for all $i = 1, \ldots, k$, as given by Equation (15). □

The total clustering coefficient is the average of the clustering coefficients over all agents, $C \equiv \frac{1}{n} \sum_{u \in \mathcal{N}} C(u)$. The clustering coefficients of the agents in ascending order of the graph in Figure 13, left, are $5/12, 5/12, 13/21, 9/10, 1, 1, 1, 1, 1, 1$, with a total clustering coefficient of $C = 0.84$. In the graph in Figure 13, right, it is $13/36, 13/28, 4/7, 1, 1, 1, 1, 1, 1$, with a total clustering of $C = 0.74$.

B.1.3. Assortativity and Nearest Neighbor Connectivity

There exists a measure of degree correlation called *average nearest neighbor connectivity* [Pastor-Satorras et al., 2001]. More precisely, the average nearest neighbor connectivity $d_{nn}(u)$ is the average degree of the neighbors of an agent with degree d_u. It is defined by

$$d_{nn}(u) \equiv \frac{1}{d_u} \sum_{v \in \mathcal{N}_u} d_v. \tag{16}$$

In a nested split graph the average nearest neighbor connectivity is determined by its degree partition.

Corollary 5. *Consider a nested split graph $G = (\mathcal{N}, \mathcal{E})$ and let $\mathcal{D} = (\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_k)$ be the* *degree partition of G. Denote by $S_{\mathcal{D}}^i = \sum_{j=1}^{i} |\mathcal{D}_{k+1-j}|$. Then for each $u \in \mathcal{D}_i$, $i = 0, \ldots, k$,*

$$d_{nn}(u) = \begin{cases} 0, & \text{if } i = 0, \\ \frac{1}{S_{\mathcal{D}}^i} \sum_{j=1}^{i} |\mathcal{D}_{k+1-j}| \left(S_{\mathcal{D}}^{k+1-j} - 1 \right), & \text{if } i = 1, \ldots, \left\lfloor \frac{k}{2} \right\rfloor, \\ \frac{1}{S_{\mathcal{D}}^{\lfloor \frac{k}{2} \rfloor + 1} - 1} \left[\sum_{j=\lfloor \frac{k}{2} \rfloor + 2}^{k} |\mathcal{D}_j| \left(S_{\mathcal{D}}^j - 1 \right) \right. & \\ \left. + \left(|\mathcal{D}_{\lfloor \frac{k}{2} \rfloor + 1}| - 1 \right) \left(S_{\mathcal{D}}^{\lfloor \frac{k}{2} \rfloor + 1} - 1 \right) + |\mathcal{D}_{\lfloor \frac{k}{2} \rfloor}| S_{\mathcal{D}}^{\lfloor \frac{k}{2} \rfloor} \right], & \text{if } i = \left\lfloor \frac{k}{2} \right\rfloor + 1, \quad k \text{ even}, \\ \frac{1}{S_{\mathcal{D}}^{\lfloor \frac{k}{2} \rfloor + 1} - 1} \left[\sum_{j=\lfloor \frac{k}{2} \rfloor + 1}^{k} |\mathcal{D}_j| \left(S_{\mathcal{D}}^j - 1 \right) \right] - 1, & \text{if } i = \left\lfloor \frac{k}{2} \right\rfloor + 1, \quad k \text{ odd}, \\ \frac{1}{S_{\mathcal{D}}^i - 1} \left[\sum_{j=\lfloor \frac{k}{2} \rfloor + 1}^{k} |\mathcal{D}_j| \left(S_{\mathcal{D}}^j - 1 \right) \right. & \\ \left. + \sum_{j=k-i+1}^{\lfloor \frac{k}{2} \rfloor} |\mathcal{D}_j| S_{\mathcal{D}}^j \right] - 1, & \text{if } i = \left\lfloor \frac{k}{2} \right\rfloor + 2, \ldots, k \end{cases} \tag{17}$$

Proof of Corollary 5. First, consider an agent $u \in \mathcal{D}_i$ with $i = 1, \ldots, \left\lfloor \frac{k}{2} \right\rfloor$ corresponding to the independent sets. We know that the number of neighbors (degree) of agent u is given by $\sum_{j=1}^{i} |\mathcal{D}_{k+1-j}|$. The neighbors of agent u are the agents in the dominating subsets with degrees given in Equation (13). Thus, the number of neighbors of the neighbors of u in the sets $\mathcal{D}_{k+1-j}$ is $\sum_{l=1}^{k+1-j} |\mathcal{D}_{k+1-l}| - 1$. Putting the above results together, we obtain for the average nearest neighbor connectivity of agent $u \in \mathcal{D}_i$, $i = 1, \ldots, \left\lfloor \frac{k}{2} \right\rfloor$, the following expression.

$$d_{nn}(u) = \frac{1}{\sum_{j=1}^{i} |\mathcal{D}_{k+1-j}|} \sum_{j=1}^{i} |\mathcal{D}_{k+1-j}| \left(\sum_{l=1}^{k+1-j} |\mathcal{D}_{k+1-l}| - 1 \right). \tag{18}$$

Next, we consider an agent u in the set $\mathcal{D}_i$ with $\left\lfloor \frac{k}{2} \right\rfloor + 2 \leq i \leq k$ corresponding to the dominating subsets. The number of neighbors of agent u is given by $\sum_{j=1}^{i} |\mathcal{D}_{k+1-j}| - 1$. The number of neighbors of an agent $v \in \mathcal{D}_j$, $\left\lfloor \frac{k}{2} \right\rfloor + 1 \leq j \leq k$ in the dominating subsets is given by $\sum_{l=1}^{j} |\mathcal{D}_{k+1-l}| - 1$. Since agent u is connected to all other agents in the dominating subsets, we can sum over all their neighborhoods with a total of $\sum_{j=\lfloor \frac{k}{2} \rfloor + 1}^{k} |\mathcal{D}_j| \left(\sum_{l=1}^{j} |\mathcal{D}_{k+1-l}| - 1 \right)$ neighbors. Note however, that we have to subtract agent u herself from this sum. Morover, the number of neighbors of an agent $w \in \mathcal{D}_j$, $1 \leq j \leq \left\lfloor \frac{k}{2} \right\rfloor$ in the independent sets is given by $\sum_{l=1}^{j} |\mathcal{D}_{k+1-l}|$. Thus, the average nearest neighbor connectivity of agent $u \in \mathcal{D}_i$, $\left\lfloor \frac{k}{2} \right\rfloor + 2 \leq i \leq k$, is given by

$$d_{nn}(u) = \frac{1}{\sum_{j=1}^{i} |\mathcal{D}_{k+1-j}| - 1} \left[\sum_{j=\lfloor \frac{k}{2} \rfloor + 1}^{k} |\mathcal{D}_j| \left(\sum_{l=1}^{j} |\mathcal{D}_{k+1-l}| - 1 \right) + \sum_{j=k-i+1}^{\lfloor \frac{k}{2} \rfloor} |\mathcal{D}_j| \sum_{l=1}^{j} |\mathcal{D}_{k+1-l}| \right] - 1. \tag{19}$$

In a similar way we can consider the cases $i = \left\lfloor \frac{k}{2} \right\rfloor + 1$ for both k even and k odd. $\square$

When the average nearest neighbor connectivity is a monotonic increasing function of the degree d, then the network is assortative, while, if it is monotonic decreasing with d, it is dissortative [Newman, 2002; Pastor-Satorras et al., 2001]. Nested split graphs are dissortative, since for $i < j$ and $d_u \in \mathcal{D}_i < d_v \in \mathcal{D}_j$ it follows that $d_{nn}(u) > d_{nn}(v)$. This is because the higher is the degree of an agent in a dominating subset, the more neighbors she has from the independent sets with low degrees, which decreases her average nearest neighbor connectivity. For example, the average nearest neighbor connectivities of the agents in the graph in Figure 13, left, in ascending order are $13/3, 13/3, 37/7, 33/5, 15/2, 15/2, 25/3, 25/3, 9, 9$ while in the graph in Figure 13, right, they are $35/9, 35/8, 34/7, 7, 7, 8, 8, 8, 17/2, 9$.

B.1.4. Characteristic Path Length

The characteristic path length is defined as the number of links in the shortest path between two agents, averaged over all pairs of agents [Watts and Strogatz, 1998]. This can be written as

$$\ell(G) \equiv \frac{1}{n(n-1)/2} \sum_{u \neq v \in G} d(u,v), \tag{20}$$

where $d(u,v)$ is the geodesic (shortest path) between agent u and agent v in $N \backslash \mathcal{D}_0$.[53] Then the characteristic path length in a nested split graph is given by the following corollary.

Corollary 6. *Consider a nested split graph $G = (\mathcal{N}, \mathcal{E})$ and let $\mathcal{D} = (\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_k)$ be the degree partition of G. Then the characteristic path length of G is given by*

$$\begin{aligned} \ell(G) \ &= \tfrac{1}{n(n-1)/2}\left[\tfrac{1}{2}\sum_{j=\lfloor \frac{k}{2}\rfloor+1}^{k}|\mathcal{D}_j|\left(\sum_{j=\lfloor \frac{k}{2}\rfloor+1}^{k}|\mathcal{D}_j|-1\right)+\sum_{j=1}^{\lfloor \frac{k}{2}\rfloor}|\mathcal{D}_j|\left(\sum_{j=1}^{\lfloor \frac{k}{2}\rfloor}|\mathcal{D}_j|-1\right)\right. \\ &\quad \left. +\sum_{l=1}^{\lfloor \frac{k}{2}\rfloor}|\mathcal{D}_l|\left(\sum_{j=k-l+1}^{k}|\mathcal{D}_j|+2\sum_{j=\lfloor \frac{k}{2}\rfloor+1}^{k-l}|\mathcal{D}_j|\right)\right]. \end{aligned} \tag{21}$$

Proof of Corollary 6. We first consider all pairs of agents in the dominating subsets. All theses agents are adjacent to each other and thus the shortest path between them has length one. Moreover, there are $\frac{1}{2}\sum_{j=\lfloor \frac{k}{2}\rfloor+1}^{k}|\mathcal{D}_j|\left(\sum_{j=\lfloor \frac{k}{2}\rfloor+1}^{k}|\mathcal{D}_j|-1\right)$ pairs of agents in the dominating subsets.

Next, we consider all pairs of agents in the independent sets. From Equation (26) we know that all of them are at a distance of two links separated from each other. Moreover, there are $\frac{1}{2}\sum_{j=1}^{\lfloor \frac{k}{2}\rfloor}|\mathcal{D}_j|\left(\sum_{j=1}^{\lfloor \frac{k}{2}\rfloor}|\mathcal{D}_j|-1\right)$ pairs of agents in which both agents stem from an independent set.

Further, consider the pairs of agents in which one agent is in the independent set $\mathcal{D}_1$ and the other in a dominating subset. Then there are $|\mathcal{D}_1||\mathcal{D}_k|$ pairs of agents with shortest path 1 and $|\mathcal{D}_1|\sum_{j=\lfloor \frac{k}{2}\rfloor+1}^{k-1}|\mathcal{D}_j|$ pairs of agents with shortest path 2. Similarly, we can consider the pairs in which one agent is in the set $\mathcal{D}_2$. Then we have $|\mathcal{D}_2|(|\mathcal{D}_k|+|\mathcal{D}_{k-1}|)$ pairs of agents with shortest path 1 and $|\mathcal{D}_2|\sum_{j=\lfloor \frac{k}{2}\rfloor+1}^{k-2}|\mathcal{D}_j|$ pairs of agents with shortest path 2. Finally, if one agent is in the set $\mathcal{D}_{\lfloor \frac{k}{2}\rfloor}$ then we have $|\mathcal{D}_{\lfloor \frac{k}{2}\rfloor}|\sum_{j=\lfloor \frac{k}{2}\rfloor+1}^{k}|\mathcal{D}_j|$ pairs of agents with distance 1 and none with distance 2, if k is even (see Figure 13, left). If k is odd (see Figure 13, right), and we have one agent is in the set $\mathcal{D}_{\lfloor \frac{k}{2}\rfloor}$ then we have $|\mathcal{D}_{\lfloor \frac{k}{2}\rfloor}|\sum_{j=\lfloor \frac{k}{2}\rfloor+2}^{k}|\mathcal{D}_j|$ pairs of agents with distance 1 and $|\mathcal{D}_{\lfloor \frac{k}{2}\rfloor}||\mathcal{D}_{\lfloor \frac{k}{2}\rfloor+1}|$ pairs with distance 2.

Therefore, the average path length $\ell(G)$ defined in Equation (20) is given by the following equation

$$\begin{aligned} \tfrac{n(n-1)}{2}\ell(G) \ &= \tfrac{1}{2}\sum_{j=\lfloor \frac{k}{2}\rfloor+1}^{k}|\mathcal{D}_j|\left(\sum_{j=\lfloor \frac{k}{2}\rfloor+1}^{k}|\mathcal{D}_j|-1\right)+2\tfrac{1}{2}\sum_{j=1}^{\lfloor \frac{k}{2}\rfloor}|\mathcal{D}_j|\left(\sum_{j=1}^{\lfloor \frac{k}{2}\rfloor}|\mathcal{D}_j|-1\right) \\ &\quad +\sum_{l=1}^{\lfloor \frac{k}{2}\rfloor}|\mathcal{D}_l|\left[\sum_{j=k-l+1}^{k}|\mathcal{D}_j|+2\sum_{j=\lfloor \frac{k}{2}\rfloor+1}^{k-l}|\mathcal{D}_j|\right]. \end{aligned} \tag{22}$$

□

[53] Note that we do not consider the isolated agents in the set $\mathcal{D}_0$ because the characteristic path length $\ell(G)$ is not defined for disconnected networks G.

Considering the graph in Figure 13, left, the characteristic path length is $\ell(G) = 22/15$ while in the graph in Figure 13, right, we get $\ell(G) = 68/45$.

Note that, by taking the inverse of the shortest path length one can introduce a related measurement, the network efficiency,[54] $\epsilon(G) \equiv \frac{1}{n(n-1)} \sum_{u \neq v \in G} \frac{1}{d(u,v)}$ that is also applicable to disconnected networks [Latora and Marchiori, 2001]. Finally, we find that in a connected nested split graph agents are at most two links separated from each other and thus these graphs are characterized by a short characteristic path length.

B.2. Centrality

In the next sections we analyze different measures of centrality in a nested split graph G. We derive the expressions for degree, closeness and betweenness centrality as a function of the degree partition of G. Finally, we show that these measures are similar in the sense that they induce the same ordering of the nodes in G based on their centrality values.

B.2.1. Degree Centrality

The degree centrality of an agent $u \in \mathcal{N}$ is given by the proportion of agents that are adjacent to u [Wasserman and Faust, 1994]. We obtain the normalized degree centrality simply by dividing the degree of agent u with the maximum degree $n-1$. This yields the following corollary.

Corollary 7. *Consider a nested split graph $G = (\mathcal{N}, \mathcal{E})$ and let $\mathcal{D} = (\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_k)$ be the degree partition of G. Then for each $u \in \mathcal{D}_i$, $i = 0, \ldots, k$, the degree centrality is given by*

$$C_d(u) = \begin{cases} \frac{1}{n-1} \sum_{j=1}^{i} |\mathcal{D}_{k+1-j}|, & \text{if } 1 \leq i \leq \left\lfloor \frac{k}{2} \right\rfloor, \\ \frac{1}{n-1} \left(\sum_{j=1}^{i} |\mathcal{D}_{k+1-j}| - 1 \right), & \text{if } \left\lfloor \frac{k}{2} \right\rfloor + 1 \leq i \leq k. \end{cases} \tag{23}$$

Proof of Corollary 7. The result follows directly from Corollary 3. □

We observe that degree centrality as well as the degree are increasing with increasing index i of the set $\mathcal{D}_i$ to which agent u belongs. Degree centralities for the graphs shown in Figure 13 can be derived from the degrees given in Section B.1.1 by dividing the degrees with $n-1$.

B.2.2. Closeness Centrality

Excluding the isolated nodes in G, closeness centrality of agent $u \in \mathcal{N} \setminus \mathcal{D}_0$ is defined as [Beauchamp, 1965; Sabidussi, 1966]:

$$C_c(u) = \frac{n-1}{\sum_{v \neq u \in G} d(u,v)}. \tag{24}$$

where $d(u,v)$ measures the shortest path between agent u and agent v in $\mathcal{N} \setminus \mathcal{D}_0$. For a nested split graph we obtain the following corollary.

[54] The network efficiency must not be confused with the efficiency of a network. The first is related to short paths in the network while the latter measures social welfare, that is, the efficient network maximizes aggregate payoff.

Corollary 8. *Consider a nested split graph $G = (\mathcal{N}, \mathcal{E})$ and let $\mathcal{D} = (\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_k)$ be the degree partition of G. Then for each $u \in \mathcal{D}_i$, $i = 0, \ldots, k$, the closeness centrality is given by*

$$C_c(u) = \begin{cases} \frac{n-1}{\sum_{j=k-i+1}^{k} |\mathcal{D}_j| + 2\sum_{j=1}^{k-i} |\mathcal{D}_j| - 2}, & \text{if } 1 \leq i \leq \left\lfloor \frac{k}{2} \right\rfloor, \\ \frac{n-1}{\sum_{j=k-i+1}^{k} |\mathcal{D}_j| + 2\sum_{j=1}^{k-i} |\mathcal{D}_j| - 1}, & \text{if } \left\lfloor \frac{k}{2} \right\rfloor + 1 \leq i \leq k. \end{cases} \tag{25}$$

Proof of Corollary 8. For both agents in the independent sets, $u \in \mathcal{D}_i$ with $1 \leq i \leq \left\lfloor \frac{k}{2} \right\rfloor$, and in the dominating subsets, $u \in \mathcal{D}_i$ with $\left\lfloor \frac{k}{2} \right\rfloor + 1 \leq i \leq k$, we can compute the length of the shortest paths as follows:

$$d(u, v) = \begin{cases} 1 & \text{for all } v \in \bigcup_{j=k-i+1}^{k} \mathcal{D}_j, \\ 2 & \text{for all } v \in \bigcup_{j=1}^{k-i} \mathcal{D}_j. \end{cases} \tag{26}$$

In order to compute the closeness centrality we have to consider all pairs of agents in the graph and compute the length of the shortest path between them, which is given in Equation (26). We obtain for any agent $u \in \mathcal{D}_i$, $i = 1, \ldots, k$, the following expression

$$C_c(u) = \begin{cases} \frac{n-1}{\sum_{j=k-i+1}^{k} |\mathcal{D}_j| + 2\sum_{j=1}^{k-i} |\mathcal{D}_j| - 2}, & \text{if } 1 \leq i \leq \left\lfloor \frac{k}{2} \right\rfloor \\ \frac{n-1}{\sum_{j=k-i+1}^{k} |\mathcal{D}_j| + 2\sum_{j=1}^{k-i} |\mathcal{D}_j| - 1}, & \text{if } \left\lfloor \frac{k}{2} \right\rfloor + 1 \leq i \leq k. \end{cases} \tag{27}$$

Note that we have subtracted 1 and 2 in the denominator, respectively, since the sums would otherwise include the contribution of agent u herself. □

We have that closeness centrality is identical for all agents in the same set. Also note that $C_c(u) = 1$ for $u \in \mathcal{D}_k$. Moreover, closeness centrality is increasing with increasing degree. The closeness centralities of the agents in descending order for the graph in Figure 13, left, are $1, 1, 9/11, 9/13, 9/14, 9/14, 9/15, 9/15, 9/16, 9/16$ while in the graph in Figure 13, right, they are $1, 9/10, 9/11, 9/14, 9/14, 9/15, 9/15, 9/15, 9/16, 9/17$.

B.2.3. Betweenness Centrality

Betweenness centrality is defined as [Freeman, 1977]

$$C_b(u) = \sum_{u \neq v \neq w \in G} \frac{g(v, u, w)}{g(v, w)}, \tag{28}$$

where $g(v, w)$ denotes the number of shortest paths from agent v to agent w and $g(v, u, w)$ counts the number of paths from agent v to agent w that pass through agent u.

The betweenness centrality for a nested split graph can be derived from its degree partition as follows.[55]

Corollary 9. *Consider a nested split graph $G = (\mathcal{N}, \mathcal{E})$ and let $\mathcal{D} = (\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_k)$ be the degree partition of G. Then $c_b(u) = 0$ if $u \in \mathcal{D}_i$, $i = 0, \ldots, \left\lfloor \frac{k}{2} \right\rfloor$ and for each $u \in \mathcal{D}_i$, $v \in \mathcal{D}_{i-1}$,*

[55] A similar result can be found in Hagberg et al. [2006].

$i = \left\lfloor \frac{k}{2} \right\rfloor + 1, \ldots, k$, *the betweenness centrality is given by*

$$C_b(u) = \begin{cases} 0 & \text{if}\ , i = \left\lfloor \frac{k}{2} \right\rfloor + 1, \quad k \quad \text{odd} \\ \frac{|\mathcal{D}_{\lfloor \frac{k}{2} \rfloor}|\left(|\mathcal{D}_{\lfloor \frac{k}{2} \rfloor}|-1\right)}{\sum_{j=\lfloor \frac{k}{2} \rfloor+1}^{k} |\mathcal{D}_j|}, & \text{if}\ , i = \left\lfloor \frac{k}{2} \right\rfloor + 1, \quad k \quad \text{even} \\ c_b(v) + \frac{|\mathcal{D}_{k-i+1}|(|\mathcal{D}_{k-i+1}|-1)}{\sum_{j=i}^{k} |\mathcal{D}_j|} + \frac{2|\mathcal{D}_{k-i+1}| \sum_{j=k-i+2}^{i-1} |\mathcal{D}_j|}{\sum_{j=i}^{k} |\mathcal{D}_j|}, & \text{if}\ \left\lfloor \frac{k}{2} \right\rfloor + 2 \leq i \leq k. \end{cases} \tag{29}$$

Proof of Corollary 9. In this proof, we follow closely Hagberg et al. [2006]. The agents in the independent sets $\mathcal{D}_i$, $0 \leq i \leq \left\lfloor \frac{k}{2} \right\rfloor$ do not lie on any shortest path between two other agents in the network and thus their betweenness centrality vanishes. For the agents in the dominating subsets we have that the betweenness centrality of the agent $u \in \mathcal{D}_{\lfloor \frac{k}{2} \rfloor+1}$ vanishes if k is odd and is given by $|\mathcal{D}_{\lfloor \frac{k}{2} \rfloor}| \left(|\mathcal{D}_{\lfloor \frac{k}{2} \rfloor}| - 1\right) / \sum_{j=\lfloor \frac{k}{2} \rfloor+1}^{k} |\mathcal{D}_j|$ if k is even. The latter result is due the shortest path between agents that are both in $\mathcal{D}_{\lfloor \frac{k}{2} \rfloor}$. Next, consider an agent $u \in \mathcal{D}_i$ and $v \in \mathcal{D}_{i-1}$, with $\left\lfloor \frac{k}{2} \right\rfloor + 2 \leq i \leq k$. Then the betweenness centrality of agent u is given by the following recursive relationship

$$C_b(v) + \frac{|\mathcal{D}_{k-i+1}| \left(|\mathcal{D}_{k-i+1}| - 1\right)}{\sum_{j=i}^{k} |\mathcal{D}_j|} + \frac{2|\mathcal{D}_{k-i+1}| \sum_{j=k-i+2}^{i-1} |\mathcal{D}_j|}{\sum_{j=i}^{k} |\mathcal{D}_j|}. \tag{30}$$

The first term in Equation (30) is due to the fact that all shortest paths through lower dominating nodes $v \in \mathcal{D}_{i-1}$ have the same length as through $u \in \mathcal{D}_i$. The second term in Equation (30) represents the contribution of paths between nodes in $\mathcal{D}_{k-i+1}$, divided by the number of shortest path passing through the agents in the dominating subsets $\mathcal{D}_j$, $i \leq j \leq k$. The third term in Equation (30) represents all path between an agent in $\mathcal{D}_{k-i+1}$ and the other being in $\mathcal{D}_j$, $k - i + 2 \leq j \leq i - 1$, divided by the number of shortest path passing through the agents in the dominating subsets $\mathcal{D}_j$, $i \leq j \leq k$. □

From Corollary 9, we find that the agents in the independent sets $\mathcal{D}_i$ with $1 \leq i \leq \left\lfloor \frac{k}{2} \right\rfloor$ have vanishing betweenness centrality. From the above equation we also observe that the betweenness centrality is increasing with degree such that the agents in $\mathcal{D}_k$ have the highest betweenness centrality, the agents in $\mathcal{D}_{k-1}$ the second highest betweenness centrality and so on. Thus, the ordering of betweenness centralities follows the degree ordering for all agents in the dominating subsets while the agents in the independent sets have vanishing betweenness centrality. For the betweenness centralities of the agents in the graph in Figure 13, left, we obtain in descending order $109/6, 109/6, 31/6, 1/2, 0, 0, 0, 0, 0, 0$ while in the graph in Figure 13, right, they are $28, 12, 6, 0, 0, 0, 0, 0, 0, 0$.

B.2.4. Eigenvector Centrality

There is a central property that holds for nested split graphs in relation to Bonacich centrality, namely that the agents with higher degree also have higher Bonacich centrality. Similar to part (i) of Proposition 1 we can give the following corollary.[56]

Corollary 10. *Let* $\mathbf{v}$ *be the eigenvector associated with the largest real eigenvalue* $\lambda_{PF}(G)$ *of the adjacency matrix* $\mathbf{A}$ *of a nested split graph* $G = (\mathcal{N}, \mathcal{E})$*. For each* $i \in \mathcal{N}$*, let* v_i *be the eigenvector*

[56] A similar result can be found in Grassi et al. [2007].

centrality of agent i. Consider a pair of agents $i, j \in \mathcal{N}$. If and only if agent i has a higher degree than agent j then i has a higher eigenvector centrality than j, i.e. $d_i > d_j \Leftrightarrow v_i > v_j$.

Proof of Corollary 10. The proof is identical to the proof of part (i) of Proposition 1. □

B.2.5. Centrality Rankings

Putting together the results for different centrality measures derived in the previous sections, we can make the following observation of the rankings of agents for different centrality measures in a nested split graph.

Corollary 11. *Consider a nested split graph $G = (\mathcal{N}, \mathcal{E})$. Let $\mathbf{C}_d$, $\mathbf{C}_c$, $\mathbf{C}_b$, $\mathbf{C}_v$ denote the vectors of degree, closeness, betweenness and eigenvector centrality in G. Then for any $l, m \in \{d, c, v\}$, $l \neq m$ and $i, j \in \mathcal{N}$ we have that $C_l(i) \geq C_l(j) \Leftrightarrow C_m(i) \geq C_m(j)$, and $C_l(i) \geq C_l(j) \Rightarrow C_b(i) \geq C_b(j)$.*

Proof of Corollary 11. The proof is a direct application of Corollaries 7, 8 9 and Proposition 1. □

If and only if an agent i has the k-th highest degree centrality then i is the agent with the k-th highest closeness and eigenvector centrality. This result also holds for Bonacich centrality (see Proposition 1). Moreover, if an agent i has the k-th highest degree centrality then she also has the k-th highest betweenness centrality and this also holds for closeness, eigenvector and Bonacich centrality, respectively. The ordering induced by degree, closeness eigenvector and Bonacich centrality coincide and these orderings also apply in a weak sense for betweenness centrality. We discuss in Section 9 that this allows us to generalize our model to various other centrality measures beyond Bonacich centrality.

C. Introducing the Cost of Forming a Link (For Online Publication)

Consider the network formation process of Definition 2 with one modification: The agent who wants to create a link needs to pay a cost c and creates the link only of it increases her payoff. The following proposition gives a bound on the linking cost c such that *link monotonicity* holds,[57] that is, marginal payoffs from forming a link are always positive.

Proposition 8. *Consider the network formation process $(G(t))_{t \in \mathbb{R}_+}$ in Definition 2. Assume that there is a cost $c \geq 0$ of creating a link for the agent who initiates that link. Further, assume that agents create links only if it increases their payoff. Then, if c is smaller than $\lambda(2-\lambda)/(2(1-\lambda)^2)$, link monotonicity holds and the emerging network will always be a nested split graph.*

[57] Link monotonicity requires that $\pi_i^*(G \oplus ij, \lambda) > \pi_i^*(G, \lambda)$ for all $i, j \in \mathcal{N}$ and $ij \notin G$ [cf. e.g. Dutta et al., 2005].

Proof of Proposition 8. For notational simplicity we drop the time index and denote by $G \equiv G(t)$. We consider the network $G \oplus ij$ obtained by adding the link $ij \notin G$. The marginal payoff from forming a link ij for agent $i \in \mathcal{N}$ is given by

$$\begin{aligned}\pi_i^*(G \oplus ij, \lambda) - \pi_i^*(G, \lambda) &= \frac{1}{2}\left(b_i(G \oplus ij, \lambda)^2 - b_i(G, \lambda)^2\right) \\ &= \frac{1}{2}\left(b_i(G \oplus ij, \lambda) - b_i(G, \lambda)\right)\left(b_i(G \oplus ij, \lambda) + b_i(G, \lambda)\right).\end{aligned}$$

Note that the Bonacich centrality of agent $i \in \mathcal{N}$ can be written as $b_i(G, \lambda) = 1 + \sum_{j \in N_i} b_j(G, \lambda)$. The change in the Bonacich centrality from forming the link ij is given by

$$\begin{aligned}b_i(G \oplus ij, \lambda) - b_i(G, \lambda) &= \underbrace{\sum_{k \in N_i \setminus \{j\}} \left(b_k(G \oplus ij, \lambda) - b_k(G, \lambda)\right)}_{>0} + \lambda b_j(G \oplus ij, \lambda) \\ &\geq \lambda b_j(G \oplus ij, \lambda) \geq \lambda \min_{k \in \mathcal{N}} b_k(G \oplus ij, \lambda).\end{aligned}$$

In the first line from above we have used the fact that the number of walks emanating at i is increasing with the addition of a link and so is the Bonacich centrality.

The smallest Bonacich centrality in a non-empty graph G (after the creation of a link the graph is always non-empty) is obtained in a path of length 2 (dyad), P_2, for which $b_i(P_2, \lambda) = \frac{1}{1-\lambda}$. Hence, we have that $b_i(G \oplus ij, \lambda) - b_i(G, \lambda) > \lambda/(1-\lambda)$ and $b_i(G \oplus ij, \lambda) + b_i(G, \lambda) > (2-\lambda)/(1-\lambda)$, so that the marginal payoff of agent i from forming a link ij is bounded from below by $\pi_i^*(G \oplus ij, \lambda) - \pi_i^*(G, \lambda) > \frac{\lambda(2-\lambda)}{2(1-\lambda)^2}$. This bound might seem crude but note that if the linking cost is higher than $\lambda(2-\lambda)/(2(1-\lambda)^2)$, the empty graph is a stable network, irrespective of how we allow agents to remove links. Hence, we find that if the cost c of a link is lower than $\lambda(2-\lambda)/(2(1-\lambda)^2)$, a link will always be formed and the networks generated in our network formation process will all be nested split graphs. □

The above proposition shows that nested split graphs can also arise even when links are costly to be formed, as long as the costs are not too large.

D. Proofs of Propositions, Corollaries and Lemmas (For Online Publication)

In this section we give the proofs of the propositions, corollaries and lemmas stated earlier in the paper.

Proof of Proposition 1. (i) A graph having a stepwise adjacency matrix is a nested split graph G. A nested split graph has a nested neighborhood structure. The neighborhood $\mathcal{N}_j$ of an agent j is contained in the neighborhood $\mathcal{N}_i$ of the next higher degree agent i with $|\mathcal{N}_i| = d_i > |\mathcal{N}_j| = d_j$ with $\mathcal{N}_j \subset \mathcal{N}_i$. For a symmetric adjacency matrix the vector of Bonacich centralities is given by $\mathbf{b}(G, \lambda) = \lambda \mathbf{A}\mathbf{b} + \mathbf{u}$, $\mathbf{u} = (1, \ldots, 1)^\top$. For agent i we get

$$b_i(G, \lambda) = \lambda \sum_{k=1}^{n} a_{ik} b_k(G, \lambda) + 1 = \lambda \sum_{k \in \mathcal{N}_i} b_k(G, \lambda) + 1, \tag{31}$$

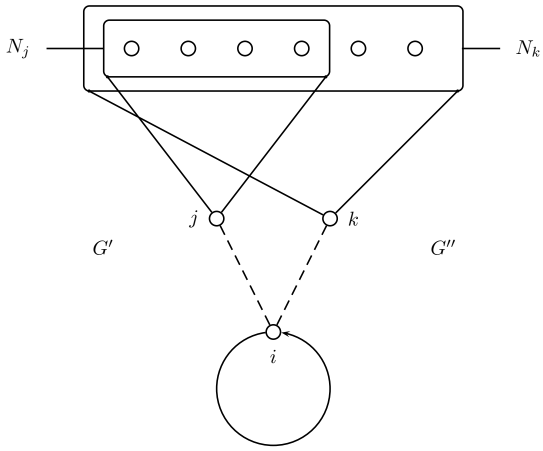


Figure 14: An illustration of the two networks G' and G'', which differ in the links ij and ik. The neighborhood $\mathcal{N}_j$ of agent j and the neighborhood $\mathcal{N}_k$ of agent k are indicated by corresponding boxes. Note that the neighborhood of agent j is contained in the neighborhood of agent k. The loop at agent i indicates a walk starting at i and coming back to i before proceeding to either agent j or k.

and similarly for agent j

$$b_j(G, \lambda) = \lambda \sum_{k \in \mathcal{N}_j} b_k(G, \lambda) + 1. \tag{32}$$

Since $\mathcal{N}_j \subset \mathcal{N}_i$ and $d_j = |\mathcal{N}_j| < |\mathcal{N}_i| = d_i$ we get

$$\frac{b_i(G, \lambda)}{b_j(G, \lambda)} = \frac{\lambda \sum_{k \in \mathcal{N}_i} b_k(G, \lambda) + 1}{\lambda \sum_{k \in \mathcal{N}_j} b_k(G, \lambda) + 1} > 1. \tag{33}$$

The inequality follows from the fact that the Bonacich centrality is nonnegative and the numerator contains the sum over the same positive numbers as the denominator plus some additional values.

Conversely, in a nested split graph we must either have $\mathcal{N}_i \subset \mathcal{N}_j$ or $\mathcal{N}_j \subset \mathcal{N}_i$. Assuming that $b_i(G, \lambda) > b_j(G, \lambda)$ we can conclude from the above equation that $\mathcal{N}_j \subset \mathcal{N}_i$ and therefore $|\mathcal{N}_i| = d_i > |\mathcal{N}_j| = d_j$. If there are l distinct degrees in G then the ordering of degrees $d_1 > d_2 > \ldots > d_l$ is equivalent to the ordering of the Bonacich centralities $b_1(G, \lambda) > b_2(G, \lambda) > \ldots > b_l(G, \lambda)$.
(ii) Consider the agents i, j and k in the nested split graph $G(t)$, such that $d_j \leq d_k$. Let G' be the graph obtained from $G(t)$ by adding the link ij and G'' be the graph obtained from $G(t)$ by adding the link ik. We want to show that the Bonacich centrality of agent i in G'' is higher than in G', that is, $b_i(G', \lambda) < b_i(G'', \lambda)$. For this purpose we count the number of walks emanating at agent i when connecting either to agent j or agent k. Since G is a nested split graph, we have that $\mathcal{N}_j \subset \mathcal{N}_k$. An illustration is given in Figure 14. We consider a walk W_l of length $l \geq 2$ starting at agent i in G'. We want to know how many such walks there are in G' and G'', respectively. For this purpose we distinguish the following cases:

(a) Assume that W_l does not contain the link ij nor the link ik. Then each such walk W_l in G' is also contained in G'', since G' and G'' differ only in the links ij and ik.
(b) Consider the graph G' and a walk W_l starting at agent i and proceeding to agent j. For each walk W_l in G' there exists a walk $\tilde{W}_l$ in G'' being identical to W_l except of proceeding from i to j it proceeds from i to k and then to the neighbor of j that is visited after j in W_l. This is always possible since the neighbors of j are also neighbors of k.
(c) Consider a walk W_l in G' that starts at i but first takes a detour returning to i before

proceeding from i to j. Using the same argument as in (ii) it follows that for each such walk W_l in G' there exists a walk of the same length in G''.

(d) Consider a walk W_l in G' that starts at agent i and at some point in its sequence of agents and links proceeds from agent j to agent i. For each such walk W_l in G' there exists a walk $\tilde{W}_l$ in G'' that is identical to W_l except that it does not proceed from a neighbor of j to j and then to i it proceeds from a neighbor of j to k and then to i.

The above cases take into account all possible walks in G' and G'' of an arbitrary length l and show that in G'' there are at least as many walks of length l starting from agent i as there are in G'.

Now consider the walks of length two, W_2, in G' starting at agent i and proceeding to agent j. Then there are $|\mathcal{N}_j|$ such walks in G'. However, there are $|\mathcal{N}_k| > |\mathcal{N}_j|$ such walks in G'' of length two that start at agent i.

The Bonacich centrality $b_i(G(t), \lambda)$ is computed by the number of all walks in $G(t)$ starting from i, where the walks of length l are weighted by their geometrically decaying factor λ^l. We have shown that for each l the number of walks in G'' is larger or equal than the number of walks in G' and for $l = 2$ it is strictly larger. Thus, the Bonacich centrality of agent i in G'' is higher than in G'. Note that all agents in a nested split graph are at most two links separated from each other (if there exists any walk between them). Thus, the agent with the highest degree is also the agent with the highest degree among the neighbors' neighbors. From this discussion we see that in a nested split graph $G(t)$ the best response of an agent i are the agents with the highest degrees in i's second-order neighborhood. □

Proof of Proposition 2. We give a proof by induction. The induction basis is trivial. We start at $t = 0$ from an empty network $G(0) = \bar{K}_n$, which has a trivial stepwise adjacency matrix (see also the Definition 5). Since there are no links present in $\bar{K}_n$ we can omit the removal of a link. Consider a small time increment $\Delta t > 0$. During that time interval a one step transition with positive probability can only involve the creation of a link by an isolated agent. All other isolated agents are best responses of this agent. The formation of the link creates a path of length one whose adjacency matrix is stepwise. This is true because we can always find a simultaneous columns and rows permutation which makes the adjacency matrix stepwise. Thus $G(\Delta t)$ has a stepwise adjacency matrix.

Next we consider the induction step of a one step transition from $G(t)$ to $G(t + \Delta t)$. By the induction hypothesis, $G(t)$ is a nested split graph with a stepwise adjacency matrix. First, we consider the creation of a link ij. Let agent j be a best response of agent i, that is $j \in \mathcal{B}_i(G(t))$. Using Proposition 1, this means that agent i must be the agent with the highest degree not already connected to j. From the stepwise adjacency matrix $\mathbf{A}(G(t))$ of $G(t)$ (see Definition 5) we find that adding the link ij to the network $G(t)$ such that j has the highest degree among all agents not already connected to i results in a matrix $\mathbf{A}(G(t) \oplus ij)$ that is stepwise. Therefore, the network $G(t) \oplus ij$ is a nested split graph.

We give an example in Figure 15. Let the agents be numbered by the rows respectively columns of the adjacency matrix. We assume that agent 4 receives a link creation opportunity. Two possible positions for the creation of a link from agent 4, either to agent 7 or to agent 10 are indicated with boxes. Since, in a stepwise matrix, the agent in the best response set has the highest degree, agent 7 is a best response of agent 4 while agent 10 is not. It further holds that agent 4 is also a best response of agent 7, since agent 4 is the agent with the highest degree not

$$
\mathbf{A}' = \begin{pmatrix}
0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\xleftarrow[G\oplus 4,10]{} \mathbf{A} = \begin{pmatrix}
0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\xrightarrow[G\oplus 47]{} \mathbf{A}'' = \begin{pmatrix}
0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
$$

Figure 15: Two possible positions for the creation of a link from agent 4, either to agent 7 (right) or to agent 10 (left), are indicated above. Agent 7 has degree 3 while agent 10 has degree 1. Creating a link to an agent with higher degree results in higher equilibrium payoffs. Thus, the best response of agent 4 is agent 7 and not agent 10.

already connected to agent 7. Finally, we observe that creating the link 47 preserves the stepwise form of the adjacency matrix (see also Definition 5).[58]

For the decay of a link a similar argument can be applied as in the preceding discussion. Disconnecting from the agent with the smallest degree decreases the Bonacich centrality and equilibrium payoffs the least, and hence, this will be the link that decays as $\zeta \to 0$. From the properties of the stepwise matrix $\mathbf{A}(G(t))$ it then follows that the matrix $\mathbf{A}(G(t) \ominus ij)$ is stepwise.

Thus, at any time $t \geq 0$ in the network formation process $(G(t))_{t \in \mathbb{R}_+}$, $G(t)$ is a nested split graph with an associated stepwise adjacency matrix $\mathbf{A}(G(t))$. Let Ψ denote the set of nested split graphs on n nodes. It can be shown that $|\Psi| = 2^{n-1}$ [Mahadev and Peled, 1995]. We thus have shown that in the limit of vanishing mistakes, $\zeta \to 0$, the network $G(t)$ is a nested split graph almost surely, which is to say that $\mathbb{P}(G(t) \in \Psi | G(0) = \bar{K}_n) = 1$ for all $t \geq 0$. □

Proof of Corollary 1. In Proposition 2 we have shown that $G(t)$ generated by $(G(t))_{t \in \mathbb{R}_+}$ is a nested split graph for all times t. In a nested split graph, any node in the connected component is directly connected to the node(s) with maximum degree. Thus, there exists a path of at most length two from any node to any other node in the connected component. It follows that $G(t)$ consists of a connected component and possible isolated nodes. □

Proof of Proposition 3. First, we show that $(G(t))_{t \in \mathbb{R}_+}$ is a Markov chain. Since the transition rate $q^{\zeta}(G, G')$ governing the transition from a network G to a network G' depends only on the current network G, the following Markov property holds

$$
\mathbb{P}\left(G(t+s) = G' | G(s) = G, \{G(u) : 0 \leq u < s\}\right) = \mathbb{P}\left(G(t+s) = G' | G(s) = G\right),
$$

for all $t \geq 0$, $s \geq 0$ and $G, G' \in \Omega$. The number of possible networks $G(t)$ is finite for any time $t \geq 0$ and the transition rates depend on the state $G(t)$ but not on the time t. Therefore, $(G(t))_{t \in \mathbb{R}_+}$ is a finite state, continuous time, homogeneous Markov chain. Further, not that the transition rates are bounded.

Next, we show that the Markov chain is irreducible. Consider two networks $G, G' \in \Omega$. $(G(t))_{t \in \mathbb{R}_+}$ is irreducible if there exists a positive probability to pass from any G to any other G' in

[58] The adjacency matrix is uniquely defined up to a permutation of its rows and columns. Applying such a permutation, we can always find an adjacency matrix which is stepwise.

Ω. This means that there exists a sequence of networks $G_1, G_2, \ldots, G_n$ such that $q^\zeta(G, G_1)q^\zeta(G_1, G_2)\cdots q^\zeta(G_n, G$ 0. We say that G' is accessible from G. For $\zeta > 0$ the logistic function in the transition rates implies that such a sequence always exists and irreducibility follows. We then have that a unique invarinat distribution μ^ζ exists.

Next, we consider the case of $\zeta = 0$. Let Ψ be the set of nested split graphs and denote by $\bar{\Psi} = \Omega \backslash \Psi$. In the following, we show that the networks in $\bar{\Psi}$ are transient. Observe that for any network $G \in \bar{\Psi}$ and $\alpha_i > 0$ there exists a positive probability that in a finite number of consecutive transitions in the Markov chain links are created and no links are removed until the complete network $K_n \in \Omega$ is reached. Let $T < \infty$ be the time when this happens starting from $G \notin \Psi$. Note that $K_n \in \Psi$, and therefore Proposition 2 implies that all networks $G(t)$, $t > T$, visited by the chain will be in Ψ. A state G is transient if $\int_0^\infty \mathbb{P}(G(t+s) = G|G(t) = G)ds < \infty$ [see e.g. Grimmett and Stirzaker, 2001, Chap. 6]. We have that $\int_0^\infty \mathbb{P}(G(t+s) = G|G(t) = G)ds = \mathbb{E}\left(\int_0^\infty \mathbb{1}_G(G(s))ds|G(t) = G\right) \leq \mathbb{E}(T) < \infty$. Therefore, all networks which are not nested split graphs are transient and they have vanishing probability in the stationary distribution, i.e. $\mu(\bar{\Psi}) = 0$.

In the following, we show that the set Ψ is a communicating class. Similar to our previous analysis, it holds that for any $G \in \Psi$ and $\alpha_i > 0$ there exists a positive probability that in all consecutive transitions in the Markov chain links are created and no links are removed until the complete network $K_n \in \Omega$ is reached. Then for $\beta_i > 0$ there exists a positive probability that from K_n only those links decay such that the network G' remains.[59] Therefore, there exists a positive probability to pass from any network G to any other network G' with positive probability, as long as $G, G' \in \Psi$. Similarly, one can show that G is accessible from G'. States G and G' in Ψ are accessible from one-another. We say that they communicate and Ψ is a communicating class.

Thus, in the case of $\zeta \to 0$, the state space Ω can be partitioned in a communicating class Ψ and a set of transient states $\bar{\Psi}$. The long run behavior of the chain is determined by the states in recurrent class Ψ and we have a unique invariant distribution with $\lim_{\zeta \to 0} \mu^\zeta(\Psi) = 1$ [see e.g. Ethier and Kurtz, 1986; Stroock, 2005]. $\square$

Before we proceed with the proof of Proposition 4, we introduce the sampled-time Markov chain $(G(t))_{t \in \mathcal{T}}$, $\mathcal{T} \equiv \{0, \Delta t, 2\Delta t, \ldots\}$, associated with the continuous time Markov chain $(G(t))_{t \in \mathbb{R}_+}$ in the limit of $\zeta \to 0$ on the same measure space $(\Omega, \mathcal{F})$ [see e.g. Gallager, 1996, Chap. 6]. In the sampled-time Markov chain $(G(t))_{t \in \mathcal{T}}$ transitions occur only at discrete times $t \in \mathcal{T}$ separated by (small) increments of size Δt. The transition probabilities of the sampled-time Markov chain $(G(t))_{t \in \mathcal{T}}$ are given by $p(G, G') = \mathbb{P}(G(t+\Delta t) = G'|G(t) = G) = q(G, G')\Delta t$ for $G \neq G'$ and $p(G, G) = \mathbb{P}(G(t+\Delta t) = G|G(t) = G) = 1 - q(G, G)\Delta t$ such that $\sum_{G' \in \Omega} \mathbb{P}(G(t+\Delta t) = G'|G(t) = G) = 1$. As the next lemma illustrates, both chains have the same long-run behavior.

Lemma 1. *The continuous time Markov chain $(G(t))_{t \in \mathbb{R}_+}$ and the sampled time Markov chain $(G(t))_{t \in \mathcal{T}}$, $\mathcal{T} \equiv \{0, \Delta t, 2\Delta t, \ldots\}$, $\Delta t \geq 0$, have the same stationary distribution μ on Ω.*

PROOF OF LEMMA 1. To see this, consider a probability measure $\mu : \Omega \to [0, 1]$. The stationary

[59] From the adjacency matrix associated with the complete network there exists a sequence of stepwise matrices in which a link to a neighbor with the smallest degree is removed such that any other stepwise matrix can be obtained.

distribution of the sampled-time Markov chain satisfies

$$\mu(G) = \sum_{G' \in \Omega} p(G', G)\mu(G') = \sum_{G' \neq G} q(G', G)\Delta t \mu(G') + (1 - q(G, G)\Delta t)\mu(G),$$

which implies the system of equations determining the stationary distribution of the continuous time Markov chain $\mu(G)q(G, G) = \sum_{G' \neq G} q(G', G)\mu(G')$, or equivalently $\mu\mathbf{Q} = 0$. □

Hence, in order to investigate the states in the support of the stationary distribution of $(G(t))_{t \in \mathbb{R}_+}$, it suffices to study the stationary distribution of the discrete time Markov chain $(G(t))_{t \in \mathcal{T}}$. Moreover, note that in the limit of $\Delta t \downarrow 0$, also the sample paths of the two chains agree [see e.g. Gallager, 1996, Chap. 6].

One can show that the sampled-time Markov chain on the nested split graphs Ψ (it is enough to require that $G(0) \in \Psi$ such that $G(t) \in \Psi$ for all $t > 0$) is irreducible and aperiodic, and hence is ergodic. Moreover, it has a primitive transition matrix $\mathbf{P}$ defined by $(\mathbf{P})_{ij} = \mathbb{P}(G(t + \Delta t) = G_j | G(t) = G_i)$ for any $G_i, G_j \in \Psi$.

Proof of Proposition 4. In the following we consider the sampled-time Markov chain $(G(t))_{t \in \mathcal{T}}$ with $\alpha \equiv \alpha_i = 1 - \beta_i$ for all $i \in N$. Due to Lemma 1, the stationary distribution of this chain is equivalent to the continuous time Markov chain of Definition 2 as ζ converges to zero. Moreover, because of ergodicity from Proposition 3, we can assume w.l.o.g. that $G(0) \in \Psi$. It then follows that $G(t) \in \Psi$ for all $t > 0$, and therefore, we restrict the state space Ω to the set of nested split graphs Ψ.

At every step $t \in \mathcal{T}$ in the sampled-time Markov chain a link is created with probability α and a link is removed with probability $1 - \alpha$. Further, we consider the complementary chain $(G'(t))_{t \in \mathcal{T}}$ on the same state space Ω where in every period t a link is created with probability $\alpha' = 1 - \alpha$ and a link is removed with probability $1 - \alpha' = \alpha$.[60] This means that a link is removed in $G'(t)$ whenever a link is created in $G(t)$ and a link is created in $G'(t)$ whenever a link is removed in $G(t)$.

As an example, consider the network G represented by the adjacency matrix $\mathbf{A}$ in Figure 15. The complement $\bar{G}$ has an adjacency matrix $\bar{\mathbf{A}}$ obtained from $\mathbf{A}$ by replacing each one element in $\mathbf{A}$ by zero and each zero element by one, except for the elements on the diagonal. Let H be the network obtained from G by adding the link 47 (setting $a_{47} = a_{74} = 1$ in $\mathbf{A}$). The probability of this link being created and thus the probability of reaching H after the process was in G is $3\alpha/n$, either by selecting one of the two nodes with degrees three or the node with degree five to create a link. Observe that this is identical to the probability of reaching the network $\bar{H}$ from $\bar{G}$ if either the two nodes with degrees seven or the node with degree four in $\bar{G}$ are selected to remove a link (with probability $\alpha' = 1 - \alpha$).

In general we can say that, for any $G_1, G_2 \in \Omega$ we have that

$$\mathbb{P}(G(t + \Delta t) = G_2 | G(t) = G_1) = \mathbb{P}\left(G'(t + \Delta t) = \bar{G}_2 | G'(t) = \bar{G}_1\right). \tag{34}$$

[60]Two nodes of $G'(t)$ are adjacent if and only if they are not adjacent in $G(t)$. Note that the complement of a nested split graph is a nested split graph as well [Mahadev and Peled, 1995]. In particular, the networks $G'(t)$ are nested split graphs in which the number of nodes in the dominating subsets corresponds to the number of nodes in the independent sets in $G(t)$ and the number of nodes in the independent sets in $G'(t)$ corresponds to the number of nodes in the dominating subsets in $G(t)$. Thus, $(G'(t))_{t \in \mathcal{T}}$ has the same state space Ω as $(G(t))_{t \in \mathcal{T}}$, namely the space Ψ consisting all unlabeled nested split graphs on n nodes.

Next consider the stationary distribution μ of $(G(t))_{t\in\mathcal{T}}$ and the corresponding transition matrix $\mathbf{P}$. Similarly, consider the stationary distribution μ' of $(G'(t))_{t\in\mathcal{T}}$ and the corresponding transition matrix $\mathbf{P}'$. Further, consider an ordering of states $G_1, G_2, \ldots$ in Ω and the transition matrix $\mathbf{P}$ with elements $(\mathbf{P})_{ij}$ giving the probability of observing G_j after the Markov chain $(G(t))_{t\in\mathcal{T}}$ was in G_i. Similarly, consider an ordering of states $\bar{G}_1, \bar{G}_2, \ldots$ in Ω and the transition matrix $\mathbf{P}'$ with elements $(\mathbf{P}')_{ij}$ giving the probability of observing $\bar{G}_j$ after the Markov chain $(G'(t))_{t\in\mathcal{T}}$ was in $\bar{G}_i$. Equation (34) implies that $\mathbf{P} = \mathbf{P}'$. Moreover, for the stationary distributions it must hold that $\mu\mathbf{P} = \mu$ and $\mu'\mathbf{P}' = \mu'$. Since $\mathbf{P}$ is primitive, $\mathbf{P}$ has a unique positive eigenvector and therefore $\mu' = \mu$. It follows that for any network $G \in \Omega$ with probability $\mu(G)$ we can take the complement $\bar{G} = G'$ and assign it the probability $\mu(G)$ to get the corresponding probability in μ', i.e. $\mu(G) = \mu'(G')$. □

Before we proceed with the proof of Proposition 5, we state two useful lemmas.

Lemma 2. *Consider the sampled-time Markov chain $(G(t))_{t\in\mathcal{T}}$, $\mathcal{T} \equiv \{0, \Delta t, 2\Delta t, \ldots\}$, $\Delta t \equiv 1/n$, with $\alpha \equiv \alpha_i = 1 - \beta_i$ for all $i \in \mathcal{N}$, $0 < \alpha \leq 1/2$, and restrict the state space Ω to the set Ψ of of all nested split graphs on n nodes. For any $0 \leq d \leq n-1$ let X denote the set of states in Ω in which there is exactly one node with degree $d+1$ and Y the set of states where there is no node with degree $d+1$. Denote by μ_X the probability of the states in X in the stationary distribution μ of $(G(t))_{t\in\mathcal{T}}$ and by μ_Y the probability of states in Y. If the number N_d of nodes with degree d in Y is $\Theta(n)$ such that $\lim_{n\to\infty} N_d/n > 0$ then $\lim_{n\to\infty} \mu_Y = 0$.*[61]

Proof of Lemma 2. Let $N(X, Y, y)$ be the expected number of times states in X occur before the process reaches Y (not counting the process as having immediately reached Y if $y \in Y$) when the process starts in y. Then the following relation holds (see Theorem 6.2.3 in Kemeny and Snell [1960] and also Ellison [2000])

$$\frac{\mu_X}{\mu_Y} = N(X, Y, y). \tag{35}$$

Let p_{YX} denote a lower bound on the probability that a state in X occurs after the process is in a state in Y and, conversely, let p_{XY} denote the probability that a state in Y occurs after the process is in a state in X. This probability is the same for all states in X, since from the properties of the Markov chain $(G(t))_{t\in\mathcal{T}}$, it follows that $p_{XY} = 2(1-\alpha)/n$, because there exist two possibilities to remove the link of the node with degree $d+1$ and the probability to select a node for link removal is $(1-\alpha)/n$. Observe that this probability vanishes for large n, and $\lim_{n\to\infty} p_{XY} = 0$. Moreover, we have that

$$\begin{aligned} N(X, Y, y) &\geq p_{YX}p_{XY} + 2p_{YX}(1-p_{XY})p_{XY} + 3p_{YX}(1-p_{XY})^2 p_{XY} + \ldots \\ &= p_{YX}p_{XY}\sum_{i=1}^{\infty} i(1-p_{XY})^{i-1} = \frac{p_{YX}}{p_{XY}}. \end{aligned}$$

The right hand side of the above inequality takes into account the fact that states in X can be reached once, twice, etc., before a state in Y is reached and assigns the corresponding probabilities to compute the expected value.

[61] By $f = \Theta(g)$ we mean that $0 < \liminf_{n\to\infty} \left|\frac{f(n)}{g(n)}\right| \leq \limsup_{n\to\infty} \left|\frac{f(n)}{g(n)}\right| < \infty$. In particular, $f = \Theta(1)$ implies that $0 < \lim_{n\to\infty} f(n) < \infty$.

By assuming that there exists a number N_d of nodes with degree d which is $\Theta(n)$, we have that $p_{YX} \geq \alpha N_d/n$ and $\lim_{n\to\infty} p_{YX} > 0$. It then follows that

$$\frac{\mu_X}{\mu_Y} = N(X,Y,y) \geq \frac{p_{YX}}{p_{XY}} = \frac{\alpha}{2(1-\alpha)} N_d \to \infty \quad \text{as } n \to \infty. \tag{36}$$

Since μ_X is a probability with $\mu_X \leq 1$, Equation (36) implies that $\lim_{n\to\infty} \mu_Y = 0$. □

Lemma 3. *Consider the sampled-time Markov chain $(G(t))_{t\in\mathcal{T}}$, $\mathcal{T} \equiv \{0, \Delta t, 2\Delta t, \ldots\}$, $\Delta t \equiv 1/n$, with $\alpha \equiv \alpha_i = 1 - \beta_i$ for all $i \in \mathcal{N}$, and state space Ψ consisting of all nested split graphs on n nodes. Then for $0 < \alpha \leq 1/2$ the asymptotic expected proportion of isolated nodes in the limit of large n is given by $P(0) = \frac{1-2\alpha}{1-\alpha}$.*

Proof of Lemma 3. We consider the expected change in the number of links $m(t)$ in $G(t)$ from t to $t+\Delta t$.[62] The number of links increases by one if any node which does not have the maximum degree $n-1$ is selected for creating a link. This happens with probability $\alpha\,(n - N_{n-1}(t))\,/n$. The number of links decreases whenever a node with degree higher than zero is selected for removing a link. This happens with probability $(1-\alpha)\,(n - N_0(t))\,/n$. Putting the above contributions together we can write for the expected change in the total number of links from t to $t + \Delta t$

$$\mathbb{E}\left(m(t+\Delta t)|\mathbf{N}(t)\right) - m(t) = \frac{\alpha}{n}\left(n - N_{n-1}(t)\right) - \frac{1-\alpha}{n}\left(n - N_0(t)\right). \tag{37}$$

Taking expectations on both sides of the above equation and denoting by $\bar{P}_t(d) \equiv \mathbb{E}\left(N_d(t)/n\right)$ we obtain

$$\mathbb{E}\left(m(t+\Delta t)\right) - \mathbb{E}\left(m(t)\right) = \alpha\left(1 - \bar{P}_t(n-1)\right) - (1-\alpha)\left(1 - \bar{P}_t(0)\right). \tag{38}$$

Let ρ denote the initial distribution of states, with $\rho_i = 1$ if $G_i = \bar{K}_n$ and zero otherwise. Further, let $\mathbf{m}$ be the column vector whose j-th coordinate, m_j, is the number of links of network $G_j \in \Omega$, and let $G_i = \bar{K}_n$. Then we can write

$$\begin{aligned}\mathbb{E}\left(m(t)\right) &= \mathbb{E}\left(m(t)|G(0)=G_i\right) = \sum_{G_j\in\Omega} \mathbb{P}\left(G(t)=G_j|G(0)=G_i\right) m_j = \sum_{G_j\in\Omega} \left(\mathbf{P}^\top\right)_{ij} m_j \\ &= \left(\mathbf{P}^\top \mathbf{m}\right)_i = \rho \mathbf{P}^\top \mathbf{m}.\end{aligned}$$

For large times t the expectation is computed over the invariant distribution μ. In particular, $\lim_{t\to\infty} \rho\mathbf{P}^\top = \mu$ and therefore $\lim_{t\to\infty} \mathbb{E}\left(m(t)\right) = \lim_{t\to\infty} \rho\mathbf{P}^\top\mathbf{m} = \mu\mathbf{m} = \lim_{t\to\infty} \mathbb{E}\left(m(t+\Delta t)\right)$. Thus, we can set the left hand side of Equation (38) to zero, in the limit of large t, and obtain a relationship between the asymptotic expected proportion of nodes of degree zero and one, respectively,

$$1 - 2\alpha = (1-\alpha)\bar{P}(0) - \alpha\bar{P}(n-1), \tag{39}$$

where we have denoted by $\bar{P}(d) = \lim_{t\to\infty} \bar{P}_t(d)$. Next, we consider the chain $(G'(t))_{t\in\mathcal{T}}$ which is constructed from $(G(t))_{t\in\mathcal{T}}$ by taking the complement of each network $G(t)$ in every period t (see also the proof of Proposition 4). In the following, denote the asymptotic expected number of links, $\lim_{t\to\infty} \mathbb{E}\left(m(t)\right)$, of $(G(t))_{t\in\mathcal{T}}$ by $\bar{m}$ and of $(G'(t))_{t\in\mathcal{T}}$ by $\bar{m}'$. By construction, we must

[62] We have that $2m(t) = \sum_{d=0}^{n-1} N_d(t) d$.

have that $\bar{m} = n(n-1)/2 - \bar{m}'$. From Proposition 4 we know that the Markov chain $(G'(t))_{t\in\mathcal{T}}$ has the same stationary distribution μ' as the chain $(G(t))_{t\in\mathcal{T}}$ for a link creation probability of $\alpha' = 1-\alpha$. For $\alpha = 1/2$ the two processes are identical and we must have that also their expected number of links are the same. This implies that for $\alpha = 1/2$, $\bar{m} = \bar{m}' = n(n-1)/4$. The only nested split graph with this number of links, for which the complement has the same number of links as the original graph, is the one in which each independent set is of size one and also each dominating subset has size one (except possibly for the set corresponding to the $\left(\lfloor \frac{k}{2} \rfloor + 1\right)$-th partition). Thus, for $\alpha = 1/2$ it must hold that $\bar{P}(0) = \bar{P}(n-1) = 1/n$.

Moreover, we know that for $\alpha < 1/2$ the expected number of maximally connected nodes (with degree $n-1$) is at most as large as the expected number for $\alpha = 1/2$, since the probability of links being created strictly decreases while the probability of links being removed increases for values of α below $1/2$ (and the probability of a maximally connected node losing a link strictly increases). Thus $\bar{P}(n-1) \leq 1/n$ for $\alpha \leq 1/2$, and for large n we can write Equation (39) as follows $1 - 2\alpha = (1-\alpha)\bar{P}(0)$. This is equivalent to

$$\bar{P}(0) = \frac{1-2\alpha}{1-\alpha}. \tag{40}$$

For $\alpha = 0$ no links are created and all nodes are isolated, that is $\bar{P}(0) = 1$, while for $\alpha = 1/2$ the asymptotic expected number of isolated nodes vanishes in the limit of large n. □

With these two lemmas in hand, let us now prove Proposition 5.

Proof of Proposition 5. For the proof of the proposition it is enough to consider the sampled-time Markov chain $(G(t))_{t\in\mathcal{T}}$ with $\alpha_i = 1-\beta_i = \alpha$ for all $i \in \mathcal{N}$. Due to Lemma 1, it has the same stationary distribution as the continuous time Markov chain of Definition 2 when ζ converges to zero. Moreover, because of ergodicity from Proposition 3, we can assume w.l.o.g. that $G(0) \in \Psi$ and $G(t) \in \Psi$ for all $t > 0$. We can then restrict the state space Ω to the nested split graphs $\Psi \subset \Omega$. We further assume w.l.o.g. that the step size is given by $\Delta t = 1/n$, which becomes arbitrarily small as n grows.

Note that $G(t)$ is completely determined by $\mathbf{N}(t)$ and vice versa. Thus it follows that $\{\mathbf{N}(t)\}_{t\in\mathcal{T}}$ is a Markov chain. Denote by $\bar{P}_t(d) \equiv \mathbb{E}\left(N_d(t)/n\right)$ the expected proportion of nodes with degree d at time t and let us denote by $\bar{P}(d) = \lim_{n\to\infty} \bar{P}_t(d)$; $\bar{P}(d)$ is determined by the invariant distribution μ in the limit of large times t. Lemma 3 shows that Equation (8) holds for $d = 0$. In the following we show by induction that, given that Equation (8) holds for $\bar{P}(d-1)$ and $\bar{P}(d)$, as n becomes large, also $\bar{P}(d+1)$ satisfies Equation (8) for all $0 \leq d < d^*$, in the limit of large n. For this purpose we consider (a) the expected number of isolated nodes $\mathbb{E}\left(N_0(t+\Delta t)|\mathbf{N}(t)\right)$ and (b) the expected number of nodes with degree $d = 1, \ldots, d^*$, $\mathbb{E}\left(N_d(t+\Delta t)|\mathbf{N}(t)\right)$ at time $t + \Delta t$, conditional on the current degree distribution $\mathbf{N}(t)$.

(a) Consider a particular network $G(t)$ in period t generated by $(G(t))_{t\in\mathbb{R}_+}$ and its associated degree distribution $\mathbf{N}(t)$. Figure 16 (left) shows an illustration of the corresponding stepwise matrix. In the following we compute the expected change of the number $N_0(t)$ of isolated nodes in $G(t)$.
The expected change of $N_0(t)$ due to the creation of a link has the following contributions. An agent with the highest degree k in $N_k(t)$ can create a link to an isolated agent and thus decreases the number of isolated agents by one. The expected change from this link is

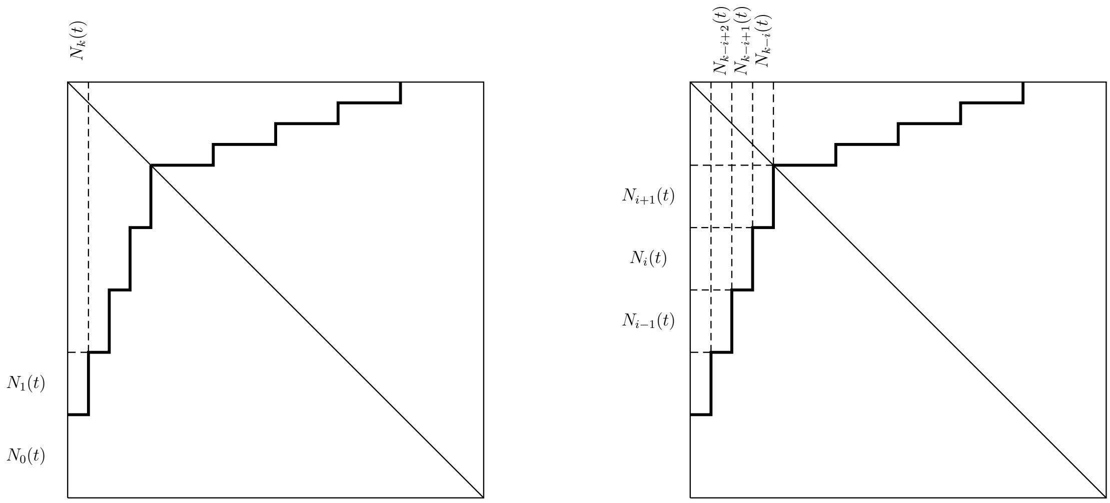


Figure 16: (Left) Representation of the stepwise matrix $\mathbf{A}$ of a nested split graph G and some selected degree partitions. The stepfunction separating the zero entries in the matrix from the one entries is shown with a thick line. (Right) Representation of the stepwise matrix $\mathbf{A}$ of a nested split graph G. The stepfunction separating the zero entries in the matrix from the one entries is shown with a thick line.

$-\alpha N_k(t)/n$. On the other hand, if an isolated agent creates a link then the expected change in the number of isolated agents is $-\alpha N_0(t)/n$.
Moreover, the removal of links can affect $N_0(t)$ if there is only one agent with maximal degree, i.e. $N_k(t) = 1$. In this case, if the agent with the highest degree removes a link, then an additional isolated agent is created yielding an expected increase in $N_0(t)$ of $(1-\alpha)N_k(t)/n$. Next, if an agent with degree one in $N_1(t)$ removes a link, then the number of isolated agents increases. Note that in a nested split graph $N_1(t) > 0$ implies that $N_k(t) = 1$ and vice versa. This gives an expected change of $N_0(t)$ given by $(1-\alpha)N_1(t)/n$.
Putting the above contributions together, the expected change in the number of isolated nodes at time $t + \Delta t$, conditional on $\mathbf{N}(t)$, is given by the following expression[63]

$$\mathbb{E}\left(N_0(t+\Delta t)|\mathbf{N}(t)\right) - N_0(t) = -\frac{\alpha}{n}\left(N_0(t) + N_k(t)\right) + \frac{1-\alpha}{n}\left(N_1(t)+1\right)\delta_{N_k(t),1}. \tag{41}$$

We can take expectations on both sides of Equation (41). For large times t the expectation is computed on the basis of the invariant distribution μ and similarly to the proof of Lemma 3, after taking expectations, we can set the left hand side of Equation (41) to zero for large times t. Note that from Lemma 3 we know that the asymptotic expected proportion $\bar{P}(0)$ of isolated nodes is $\Theta(1)$, for n large. Thus we can apply the result of Lemma 2 which tells us that the networks in which there does not exist a node with degree one have vanishing probability in μ for large n. Since the existence of a node with degree one implies that $N_k(t) = 1$, in the limit of large n we can set $\delta_{N_k(t),1} = 1$. We then obtain from Equation (41)

$$\bar{P}(1) = \frac{\alpha}{1-\alpha}\bar{P}(0). \tag{42}$$

This shows that also $\bar{P}(1)$ satisfies Equation (8). Together with Lemma 3 this proves the induction basis.

[63] $\delta_{i,j}$ denotes the usual Kronecker delta which is 1 if $i = j$ and 0 otherwise.

(b) We give a proof by induction on the number $N_d(t)$ of nodes with degree $0 < d < d^*$ in a network $G(t)$ in the support of the stationary distribution μ. In the following, we compute the expected change in $N_d(t)$ due to the creation or the removal of a link. An illustration can be found in Figure 16 (right).

Let us investigate the creation of a link. With probability α/n a link is created from an agent in $N_{k-d}(t)$ to an agent in $N_d(t)$. This yields a contribution to the expected change of $N_d(t)$ of $-\alpha N_{k-d}(t)/n$. If a link is created from an agent in $N_{k-d+1}(t)$ to an agent in $N_d(t)$ then the expected change is α/n, if $N_{k-d+1}(t)$ contains only a single agent. Similarly, if a link is created from an agent in $N_{d-1}(t)$ to an agent in $N_d(t)$ then the expected change of $N_d(t)$ is $\alpha N_{d-1}(t)/n$, if $N_{k-d+1}(t) = 1$. Moreover, if an agent in $N_d(t)$ is selected for link creation then we get an expected decrease of $-\alpha N_d(t)/n$.

Now we consider the removal of a link. If a link is removed from the agent in $N_{k-d+1}(t)$ to an agent in $N_d(t)$ then the expected change of $N_d(t)$ is $-(1-\alpha)N_{k-d+1}(t)/n$. If a link is removed from an agent in $N_{k-d}(t)$ to an agent in $N_{d+1}(t)$ then the expected increase of $N_d(t)$ is $(1-\alpha)/n$, if $N_{k-d}(t) = 1$. Moreover, if an agent in $N_{d+1}(t)$ is selected for removing a link, then we get an expected increase of $(1-\alpha)N_{i+d}(t)/n$, if $N_{k-d}(t) = 1$. Finally, if an agent in $N_d(t)$ is selected for removing a link, then we get an expected change of $-(1-\alpha)N_d(t)/n$.

Putting the above contributions together, the expected change in $N_d(t)$ is given by

$$\begin{aligned}\mathbb{E}\left(N_d(t+\Delta t)|\mathbf{N}(t)\right) - N_d(t) = & \frac{\alpha}{n}\left(-N_d(t) + \left(N_{d-1}(t)+1\right)\delta_{N_{k-d+1}(t),1} - N_{k-d}(t)\right) \\ & + \frac{1-\alpha}{n}\left(-N_d(t) + \left(N_{d+1}(t)+1\right)\delta_{N_{k-d}(t),1} - N_{k-d+1}(t)\right).\end{aligned} \tag{43}$$

We can take expectations on both sides of Equation (43) and similarly to part (a) of this proof we can set the left-hand-side of Equation (43) as t becomes large. For large times t the above expectation is computed on the basis of the invariant distribution μ. By the induction assumption, the asymptotic expected proportion $\bar{P}(d-1)$ of nodes with degree $d-1$ is $\Theta(1)$ in the limit of large n (as follows from Equation (8)). Thus we can apply Lemma 2 and neglect the networks in which there does not exist a node with degree d since they have vanishing probability in μ for large n. Similarly, we know from the induction assumption that the asymptotic proportion $\bar{P}(d)$ of nodes with degree d is $\Theta(1)$ and, by virtue of Lemma 2, we know that the networks in which there does not exist a node with degree $d+1$ have vanishing probability in μ for large n. Thus, in the limit of large n we can set $\delta_{N_{k-d+1}(t),1} = \delta_{N_{k-d}(t),1} = 1$, since the existence of nodes with degrees d and $d+1$ implies that $N_{k-d+1}(t) = N_{k-d}(t) = 1$ in the limit of large t and n. Therefore, we get from Equation (43) the following relationship

$$\bar{P}(d+1) = \frac{1}{1-\alpha}\bar{P}(d) - \frac{\alpha}{1-\alpha}\bar{P}(d-1). \tag{44}$$

Inserting the expressions for $\bar{P}(d-1)$ and $\bar{P}(d)$ from Equation (8) into Equation (44) yields

$$\bar{P}(d+1) = \frac{1}{1-\alpha}\frac{1-2\alpha}{1-\alpha}\left(\frac{\alpha}{1-\alpha}\right)^d - \frac{\alpha}{1-\alpha}\frac{1-2\alpha}{1-\alpha}\left(\frac{\alpha}{1-\alpha}\right)^{d-1} = \frac{1-2\alpha}{1-\alpha}\left(\frac{\alpha}{1-\alpha}\right)^{d+1}$$

Thus, Equation (8) also holds for $\bar{P}(d+1)$. This proves the induction step.

Finally, we have that the degree distribution must be normalized to one, i.e. $\sum_{d=0}^{n-1} \bar{P}(d) = 1$. We know that the number of agents in the dominating subsets with degrees larger than d^* is d^* (since each set contains only one node and there are d^* such sets).[64] Adding this to the number of agents in the independent sets with degree $d = 0, \ldots, d^*$ yields $n \sum_{d=0}^{d^*} \bar{P}(d) + d^* = n$. Further, inserting Equation (8) we can derive the number d^* of independent sets as a function of n and α

$$d^*(n, \alpha) = \frac{\ln\left(\frac{2(1-\alpha)}{(1-2\alpha)n}\right)}{\ln\left(\frac{\alpha}{1-\alpha}\right)}. \tag{45}$$

d^* is a monotonic decreasing function of n for a fixed value of α. Conversely, for a fixed value of n we get the limits $\lim_{\alpha\to 0} d^* = 0$ and $\lim_{\alpha \to 1/2} d^* = n/2$.

We finish the proof with the following observation, showing that the empirical degree distribution concentrates around its expected value in the limit of large n when $\Delta t = 1/n$. More precisely, for any $\varepsilon > 0$ we have that $\mathbb{P}\left(|P_t(d) - \mathbb{E}\left(P_t(d)\right)| \geq \varepsilon\right) \leq 2e^{-\frac{\varepsilon^2 n^2 \Delta t}{8t}}$. To see this, let us define the following random variable $Y_d(s) \equiv \mathbb{E}\left(N_d(t)|\mathbf{N}(s)\right)$, $s \in \mathcal{T}$. Since $\{\mathbf{N}(t), t \in \mathcal{T}\}$ is a Markov chain, the sequence $\{Y_d(s), s \in \mathcal{T}, s \leq t\}$ is a Martingale with respect to $\{\mathbf{N}(t), t \in \mathcal{T}\}$.[65] Moreover, the change in the number of nodes with degree d per period t is bounded by two, i.e. $|N_d(t) - N_d(t - \Delta t)| \leq 2$, since at most one link is added or removed in every period t and this can change the degrees of at most two nodes. Therefore, we can apply Hoeffding's inequality (see e.g. Theorem 3, Section 12.2 in Grimmett and Stirzaker [2001]), which states that for any $0 < s \leq t$ with $|Y(s) - Y(s - \Delta t)| \leq c$ and any $\varepsilon > 0$, $\mathbb{P}\left(|Y(t) - Y(0)| > \varepsilon\right) \leq 2e^{-\frac{\varepsilon^2 \Delta t}{2tc^2}}$. With $c = 2$, $Y(t) = \mathbb{E}\left(N_d(t)|\mathbf{N}(t)\right) = N_d(t)$, $Y(0) = \mathbb{E}\left(N_d(t)|\mathbf{N}(0)\right) = \mathbb{E}\left(N_d(t)\right)$ and $\Delta t = 1/n$ it then follows that

$$\mathbb{P}\left(\left|\frac{N_d(t)}{n} - \mathbb{E}\left(\frac{N_d(t)}{n}\right)\right| \geq \varepsilon\right) = \mathbb{P}\left(|N_d(t) - \mathbb{E}\left(N_d(t)\right)| \geq n\varepsilon\right) \leq 2e^{-\frac{\varepsilon^2 n^2 \Delta t}{8t}} = 2e^{-\frac{\varepsilon^2 n}{8t}} \to 0, \tag{46}$$

as $n \to \infty$. This implies that the empirical proportion $N_d(t)/n$ of nodes with degree d converges in probability to its expected value $\mathbb{E}\left(N_d(t)/n\right)$, as n becomes large.

Since $(\Omega, \mathcal{F}, \mathbb{P})$ is a discrete probability space, this also implies convergence almost surely. To see this, let $\mathcal{A}_n \equiv \left\{G \in \Omega : \left|\frac{N_d(t)}{n} - \mathbb{E}\left(\frac{N_d(t)}{n}\right)\right| \geq \varepsilon\right\}$. By Equation (46) we have that $\lim_{n\to\infty} \mathbb{P}_t(\mathcal{A}_n) = 0$. Then there exists and $n_0 \in \mathbb{N}$ such that $\mathbb{P}_t(\mathcal{A}_n) < \mathbb{P}_t(G)$ for all $G \in \Omega$ with $\mathbb{P}_t(G) > 0$ and $n > n_0$. Hence, for all $n > n_0$ we have that $\{G \in \Omega : \mathbb{P}_t(G) > 0\} \notin \mathcal{A}_n$, $\{G \in \Omega : \mathbb{P}_t(G) > 0\} \cap \mathcal{A}_n = \emptyset$ and therefore $\mathbb{P}_t\left(\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \mathcal{A}_m\right) = 0$. □

Proof of Corollary 2. The results follows directly from the functional form of $d^*(n, \alpha)$ in Proposition 5. □

Proof of Corollary 3. See Theorem 1.2.4 in Mahadev and Peled [1995]. □

[64]Note that since networks in which there does not exist a node with degree $0 \leq d \leq d^*$ in the corresponding independent set can be neglected, the structure of nested split graphs implies that all dominating subsets have size one.

[65]We have that $\mathbb{E}\left(Y_d(s)|\mathbf{N}(s - \Delta t)\right) = \mathbb{E}\left(\mathbb{E}\left(N_d(t)|\mathbf{N}(s)\right)|\mathbf{N}(s - \Delta t)\right) = \mathbb{E}\left(N_d(t)|\mathbf{N}(s - \Delta t)\right) = Y_d(s - \Delta t)$. Further, one can show that the first and second moments of $\{Y_d(s), s \leq t\}$ are bounded. Thus, $\{Y_d(s), s \leq t\}$ is a Martingale with respect to $\{\mathbf{N}(t)\}$ for $s, t \in \mathcal{T}$ [see e.g. Grimmett and Stirzaker, 2001, Chap. 12].