

Buechel, Berno; Buskens, Vincent

**Working Paper**

## The dynamics of closeness and betweenness

Working Papers, No. 398

**Provided in Cooperation with:**

Center for Mathematical Economics (IMW), Bielefeld University

Suggested Citation: Buechel, Berno; Buskens, Vincent (2008) : The dynamics of closeness and betweenness, Working Papers, No. 398, Bielefeld University, Institute of Mathematical Economics (IMW), Bielefeld,  
<http://nbn-resolving.de/urn:nbn:de:hbz:361-13483>

This Version is available at:

<http://hdl.handle.net/10419/43818>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

Working Papers

Institute of  
Mathematical  
Economics

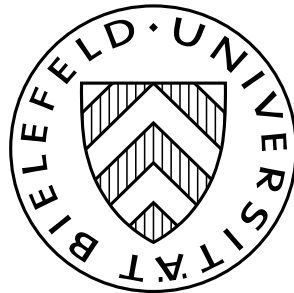
398

July 2008

## The Dynamics of Closeness and Betweenness

---

Berno Buechel and Vincent Buskens



IMW · Bielefeld University  
Postfach 100131  
33501 Bielefeld · Germany



email: [imw@wiwi.uni-bielefeld.de](mailto:imw@wiwi.uni-bielefeld.de)  
<http://www.wiwi.uni-bielefeld.de/~imw/Papers/showpaper.php?398>  
ISSN: 0931-6558

# The Dynamics of Closeness and Betweenness<sup>★</sup>

Berno Buechel<sup>a</sup> & Vincent Buskens<sup>b</sup>

<sup>a</sup>*Institute of Mathematical Economics, Bielefeld University, Germany.\**

<sup>b</sup>*Department of Sociology, Faculty of Social Sciences, Utrecht University, The Netherlands.*

---

## Abstract

Although both betweenness and closeness centrality are claimed to be important for the effectiveness of someone's network position, it has not been explicitly studied which networks emerge if actors follow incentives for these two positional advantages. We propose such a model and observe that network dynamics differ considerably in a scenario with either betweenness or closeness incentives compared to a scenario in which closeness and betweenness incentives are combined. Considering social consequences, we find low clustering when actors strive for either type of centrality. Surprisingly, actors striving for closeness are likely to reach networks with relatively low closeness and high betweenness, while this is the other way round for actors striving for betweenness. This shows that in both situations the network formation process implies a social dilemma in which the social optimum is not reached by individual optimizing.

*Key words:* Closeness Centrality, Betweenness Centrality, Actor Utility, Network Dynamics, Social Dilemma

---

---

<sup>★</sup> This paper has benefitted from audiences in Utrecht, Bielefeld, Paris, Bern, and at Sunbelt XXVIII social network conference. Buskens contribution is part of the Utrecht University High-Potential Program "Dynamics of Cooperation, Networks, and Institutions." Buechel gratefully acknowledges support of the IRTG "Economic Behavior and Interaction Models" funded by the German Science Foundation.

\* Berno Buechel, IMW, Bielefeld University, Postfach 100131, 33501 Bielefeld, Germany, phone: +49 521 106 4903, fax: +49 521 106 2997.

*Email addresses:* [berno@wiwi.uni-bielefeld.de](mailto:berno@wiwi.uni-bielefeld.de), [v.buskens@uu.nl](mailto:v.buskens@uu.nl) (Vincent Buskens).

## 1 Introduction

The importance of social networks has meanwhile been acknowledged far beyond the field of sociology. Many people are aware that social network positions are important for applicants getting a job (Granovetter, 1995), facilitating companies' cooperation (Powell et al., 1996) and success (Uzzi, 1996), and also to “Find emergent leaders in fast growing companies [...] Determine influential journalists and analysts [...] Reveal key players [...] Reveal opinion leaders” (Krebs, 2008), and so on. This awareness may affect the way individuals take decisions about social relationships. A German news magazine recently asked in the lead article whether goal-oriented choice of contacts may undermine typical friendship relations (Dillig, 2008).

Models of dynamic networks have long not considered aspects of agency. Random graph models (starting from Erdős and Renyi, 1960, until Watts, 1999, Barabási and Albert, 1999) define probabilistic processes that are able to reconstruct different patterns of empirically observed social networks. Recently, there has been much more focus on analyzing network dynamics theoretically as well as empirically in a more incentive oriented manner by relating changes back to individual propensities to form and sever ties (see Goyal, 2007, Snijders, 1996, Snijders, 2001).

### 1.1 Modeling Centrality-Oriented Network Dynamics

If the goal of studying network dynamics is to identify the “underlying mechanisms that induce network change” (Doreian and Stokman, 2003), then the goal for this paper is to uncover two such mechanisms that are both related to centrality in social networks. Three motivations justify such a model. First, it complements the theory of centrality that originally measures the effect of network positions on individual opportunities, but not the effect of individual behavior on network structure (see, e.g., Wasserman and Faust, 1994). If network dynamics show that advantageous network positions are likely *not* to be stable, effects of advantageous network positions on individual opportunities might be smaller than expected. The reason is that the advantages of the network positions can only be exploited for a short time. Second, the centrality indices are based on network statistics that are relevant in many different applications - from ancient marriages (Padgett and Ansell, 1993) to R&D collaborations (Walker et al., 1997). Third, there is empirical support for centrality being beneficial, e.g., Song et al. (2007) find that the centrality of a work unit has a positive impact on its creativity. Moreover, the claim that centrality is beneficial becomes more and more popular in the practice of business consulting, e.g., Leader Values (2008) invite to “study ways that

Leaders can make better use of networks” (other examples can be found in Krebs, 2008, or Weidner, 2008).

To cover the third aspect it is worthwhile to incorporate a centrality index that is well known. We chose two indices based on Freeman (1979): closeness centrality and betweenness centrality. In addition, we consider degree centrality because this measures the number of relations an actors maintains, which is one of the main sources of costs in establishing one’s network. By considering these three centrality indices, we cover the three most studied types of centrality measures according the typology of Borgatti and Everett (2006). Thirty years after Freeman made this classification of centrality measures, it seems a timely contribution, to integrate them in a model of network dynamics.

Choosing benefits for closeness and betweenness and costs for establishing own relations implies that our focus is not on the benefits that are derived from network closure and dense neighborhoods. Thus, we neglect these network features that are known to be beneficial, for example, for trust (e.g., Coleman, 1988, Buskens, 2002), cooperation (e.g., Raub and Weesie, 1990), and fine-grained information transfer (e.g., Krackhardt, 1992). Rather we take Granovetter’s weak tie argument as a starting point: it is the non-redundant ties that provide access to new sources of information. Burt (1992) further elaborates and strengthens the argument. Two alters that are themselves not linked not only provide access to diverse information, but also create control benefits for ego. Both closeness and betweenness increase more by connecting to distant others than to others who are already relatively close.

Together with Burt (1992), Hummon (2000), and Doreian (2006), we assume that actors evaluate the consequences of network ties and take decision according to their goals. Or as Burt (1992, p. 39) puts it: “The task for a strategic player building an efficient-effective network is to focus resources on the maintenance of bridge ties.” Accordingly, benefits of the network structure are compared to the costs of link maintenance. However, the statement suggests that the network structure is at an actor’s discretion, which is clearly not true. An actor’s network position not only depends on his linking decisions, but also on the decisions of the other actors. So rather than choosing the best network, actors interact strategically - an aspect that is best covered by game theory. Jackson and Wolinsky (1996) introduce an appropriate framework for strategic network formation. Different models build on this framework (Jackson, 2004 and Goyal, 2007 provide lists of examples), but sociological ideas were only introduced to a limited extent. Buskens and Van de Rijt (2008) study the dynamics of “structural holes.” Hummon (2000) and Doreian (2006) study the dynamics of the “connections model,” originally introduced by Jackson and Wolinsky (1996). Covering incentives for short paths, the connections model induces similar networks as when actors strive for closeness (as shown in Buechel, 2008).

What has not been done in the literature is to *contrast* and *combine* the dynamics of “closeness-type” incentives to the dynamics of “betweenness-type” incentives. There is hardly any research on the interplay between different types of incentives to predict network formation processes, although it is likely that multiple incentives are important simultaneously. For example, the Medici’s position in the marriage network was important for their trading abilities (see Padgett and Ansell, 1993). Here betweenness plays a major role, but for actors with low betweenness, it was important to be at least close to the actors with high betweenness.

## 1.2 Outline

Based on these considerations we introduce a model to examine the following research questions: What is the influence of centrality incentives on the structure of social networks? How do the dynamics of closeness differ from the dynamics of betweenness and what happens if both centrality incentives matter?

By formal derivations and computer simulation we derive properties of stable networks for pure closeness incentives, pure betweenness incentives as well as for differently weighted mixed incentives. We observe and explain why the dynamics of the mixed cases are quite distinct from the cases with pure closeness or betweenness incentives. Finally, we examine social consequences of centrality-oriented linking behavior. As expected clustering is low under all incentives we consider here. Moreover, it is not likely that actors striving for pure closeness or pure betweenness benefits reach a social optimum. In the case of closeness dynamics this leads to underconnectedness because individual actors do not want to bear the costs of bringing others close to each other. Under betweenness dynamics, individual actors also want to reap betweenness benefits that others have, but sharing these benefits over more ties is less efficient and leads to overconnectedness.

The next two sections introduce the model and some first results. Section 4 contrasts the closeness dynamics with betweenness dynamics and shows the added value of combining the two. Section 5 addresses the social consequences of individual behavior. Section 6 concludes and discusses some implications of the findings.

## 2 Model

### 2.1 Actors and Networks

We consider a finite set of actors  $N$  with typical elements  $i$  or  $j$  and size  $n \geq 3$ . The bilateral relationships among these actors are modeled as an undirected (and dichotomous) network. Let  $G$  be the set of all those networks and  $g$  a typical element. With  $ij \in g$  we denote the presence of the tie between actors  $i$  and  $j$  in  $g$ .

The distance  $d_{ij}(g)$  of two actors  $i$  and  $j$  is defined as the number of links that is minimally needed to go from actor  $i$  to  $j$  in  $g$ . Neighbors have distance 1; neighbors of neighbors that are not directly connected are at distance 2; and pairs that cannot reach each other via any number of other actors are defined to have distance  $M$ , a number larger than any possible actual distance in a network (when in need of a specific value we use  $M = n$ ).

The degree  $l_i(g)$  of an actor  $i$  is the number of links he maintains in network  $g$ . Degree can be considered as a measure of centrality (Freeman, 1979). But besides the beneficial aspects of many ties, there are also costs (time, effort, etc.) involved. We assume that the costs of maintaining relationships exceed those benefits that are restricted to direct contacts.<sup>1</sup> This means that maintaining links is costly.<sup>2</sup>

### 2.2 Closeness and Betweenness

The idea of closeness reaches back to the origins of social network analysis. An actor is considered as “central” in a social network, if his distance to other actors is small (Sabidussi, 1966). Dekker et al. (2003) argue that closeness increases accuracy of information. Song et al. (2007) provide empirical evidence for the importance of closeness for the knowledge processing of organizational units. Moreover, in the study of Powell et al. (1996) experienced firms are likely to occupy positions with high closeness. Closeness was formalized by Linton Freeman (Freeman, 1979). Freeman uses the inverse average distance of an actor to all actors in the network to measure closeness ( $\frac{n-1}{\sum_{j \in N} d_{ij}(g)}$ ). As argued in Buechel (2008), it is equally reasonable to operationalize closeness

---

<sup>1</sup> Without this assumption every actor wants to be directly linked to every other actor, independently of any other benefits.

<sup>2</sup> We only consider costs for tie maintenance and do not take into account specific costs for creating or deleting ties.

as the reverse average distance  $(-\frac{\sum_{j \in N} d_{ij}(g)}{n-1})$ . The advantage of the latter definition is that any change in closeness is proportional to a change in average distances (as also argued in Valente and Foreman, 1998). Usually closeness is not defined for actors that are not connected via any number of others. We extend closeness to all undirected networks by defining the distance for these pairs of actors to be  $M$ . In this paper, we use the normalized version of the reverse average distance. Closeness of an actor  $i$  in network  $g$  is then equal to

$$CLOSE_i(g) = \frac{M}{M-1} - \frac{\sum_{j \in N} d_{ij}(g)}{(M-1)(n-1)}. \quad (1)$$

$CLOSE_i = 0$  for isolates, while  $CLOSE_i = 1$  for an actor who is directly connected to all others in the network. As examined in Buechel (2008), this choice of operationalization (as opposed to Freeman-Closeness) is not crucial: in the models we study here, the two operationalizations lead to very similar results.

Some actors exhibit a mediating role between other actors, which can be beneficial for them. Burt (1992) emphasizes this idea by the term “tertius gaudens.” To measure the brokerage role of a certain actor he not only proposes some new measures, but also employs betweenness centrality (see Burt, 2002). Betweenness was introduced by Freeman (1979) and was shown to be beneficial in many studies thereafter (e.g., Song et al., 2007).

The betweenness of an actor  $i$  is proportional to the number of pairs  $j$  and  $k$  for whom  $i$  lays on the shortest path (also called “geodesic”). If  $j$  and  $k$  have more than one geodesics, the fraction of shortest paths going through  $i$  is used. Formally,

$$BETWEEN_i(g) = \frac{2}{(n-1)(n-2)} \sum_{j < k (j \neq i, k \neq i)} \frac{\tau_{jk}^i(g)}{\tau_{jk}(g)}, \quad (2)$$

where  $\tau_{jk}(g)$  is the number of geodesics between  $j$  and  $k$ , and  $\tau_{jk}^i(g)$  indicates the number of shortest paths between  $j$  and  $k$  that go through  $i$ ; the fraction  $\frac{\tau_{jk}^i(g)}{\tau_{jk}(g)}$  is replaced by zero, when  $\tau_{jk}(g) = 0$ . The constant before the fraction normalizes betweenness to be between zero (an actor is on no shortest path between two other actors) and one (the center in a star network).

### 2.3 Utility Function and Actor Behavior

We assume that closeness and betweenness are the benefits derived from the network structure, while direct links are costly. Let  $c > 0$  be the costs of



one link and  $\lambda \in [0, 1]$  is the relative importance (weight) of closeness versus betweenness benefits. In this “centrality model” we can represent the behavior for any actor  $i$  by the following utility function

$$u_i(g) = (1 - \lambda)CLOSE_i(g) + \lambda BETWEEN_i(g) - cDEGREE_i(g). \quad (3)$$

We analyze the model for all possible parameter combinations, as they represent different contexts including high costs and low costs  $c$  for maintaining ties as well as pure closeness incentives ( $\lambda = 0$ ), pure betweenness ( $\lambda = 1$ ), and both closeness and betweenness being important ( $0.1 < \lambda < 0.9$ ).

Four behavioral assumptions are the basis of the centrality model:

- A0 *The actors take linking decisions based on their centrality only, where closeness and betweenness are beneficial while degree is costly.*
- A1 *The utility of an actor is linear in closeness, betweenness, and degree.* This assumption implies that the effect of a change in one centrality measure are independent of the level of other centrality measures. For example, the costs of a link are independent of the number of links an actor already has and independent of his closeness and betweenness. This assumption is standard, but clearly restrictive.<sup>3</sup>
- A2 *Actors are homogeneous in respect to preferences.* It is an interesting question to ask, how networks evolve when actors differ in their preferences (see, e.g., Galeotti et al., 1992). But since applications of our model are very different in nature, we put emphasis on the different contexts that influence everybody’s choice, not on the difference between actors (as also argued in Burger and Buskens, 2006).
- A3 *The actors take linking decisions in a myopic way.* This means that actors consider the consequences of their actions on the current network structure, but do not anticipate the potential reactions of others.

### 3 Basic Results

To study which networks are likely to emerge for different incentives under the assumptions specified above, we employ three complementary methods: analytic analysis, enumeration, and simulation. In the following we introduce each method and show the results for our model.

---

<sup>3</sup> Goyal and Joshi (2006), Kamphorst (2007) and Buechel (2008) study models where the assumption is partially relaxed (by allowing for increasing and decreasing marginal returns).

### 3.1 Analytic Analysis

In order to find the networks that are likely to emerge, the first step is to exclude all those networks in which individual actors have incentives and possibilities to change the network (“stability”). Jackson and Wolinsky (1996) proposed such a condition that takes into account that, typically for social networks, the establishment of a relationship needs the agreement of both actors involved, while the dissolution can be done unilaterally. Accordingly, we will not consider a network as stable if (i) there is an actor who wants to cut a link or (ii) there are two actors who want to form a link. Formally, let  $g \cup ij$  (respectively  $g \setminus ij$ ) be the network obtained when the link between actor  $i$  and actor  $j$  is added (respectively, removed) to (from) the network  $g$ .

A network  $g$  is **(pairwise) stable** (PS) if

- (i)  $\forall ij \in g, \quad u_i(g) \geq u_i(g \setminus ij) \text{ and } u_j(g) \geq u_j(g \setminus ij) \quad \text{and}$
- (ii)  $\forall ij \notin g, \quad u_i(g \cup ij) > u_i(g) \Rightarrow u_j(g \cup ij) < u_j(g).$

The stability analysis is used to characterize the stable networks by properties they must or must not satisfy. Moreover, it is straightforward to analyze for which parameter settings a network is stable and not stable.

Let us illustrate this for the following prominent examples: The complete network in which every possible link is present, the empty network (or null network) in which no link is present, the circle network (sometimes called “ring” or “cycle”), and two examples of complete bipartite networks. A complete bipartite network consists of two groups of actors, where all the links across groups are present and there are no links within the groups.

- In the empty network, adding a link only increases closeness for the actors involved, while their betweenness remains zero. Therefore, the empty network is stable if and only if the linking costs are larger than the benefit of getting close to this one other actor.
- Similarly, in the complete network, removing a link decreases the closeness of the actors involved, while their betweenness remains zero. Therefore, the complete network is stable if and only if the linking costs are smaller than the loss related to getting further away from this one other actor.
- In the circle network, each additional link provides a significant amount of both closeness and betweenness benefits. Removing a tie also reduces both betweenness as well as closeness for both actors involved. So rather than dissolving a tie, two actors are willing to form an additional one, across the circle even if linking costs are relatively high. So the circle network can be expected to be stable only for relatively high linking costs.
- In the balanced (both groups have the same size, see Buskens and Van de Rijt, 2008) complete bipartite network, all actors have a considerable

amount of betweenness as well as high closeness, because distance is at most 2 and each pair that is at distance 2 is mediated by all the actors in the other group. Adding a tie is not very beneficial in terms of closeness and betweenness, while the loss of removing a tie is a bit larger. Consequently, if ties are rather cheap, but not too cheap these networks can be stable.

- In the star network (complete bipartite network with only one actor in one group) the center  $k$  has maximal closeness, maximal betweenness ( $CLOSE_k(g) = BETWEEN_k(g) = 1$ ), but also maximal costs. Once this network is reached, it will hardly be left, because the creation of a tie hardly increases closeness and betweenness for the peripheral actors, while a dissolution can be very harmful.

These considerations are the basis for the formal results that are illustrated in figure 1 (see also proposition 1 and its proof is in the appendix). Figure 1 depicts the parameter space with  $\lambda$  on the horizontal axis (on the left boundary only closeness matter, on the right boundary only betweenness benefits matter) and linking costs  $c$  on the vertical axis.

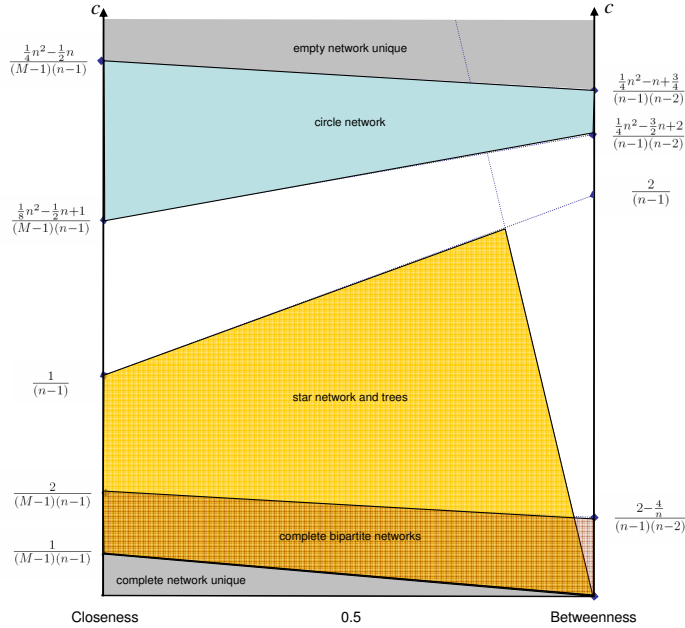


Fig. 1. “Parameter map” with stability for some prominent networks.

It is intuitive that the empty network is the uniquely stable network for very high costs, because no link justifies its costs. The complete network is uniquely stable for sufficiently low costs as long as there are some incentives for closeness, because in the complete network betweenness is also zero for everyone (see also proposition 2 in the appendix). Above the upper bound for the stability of the complete network there is (possibly) a multitude of stable networks. Among them is the star network, which is stable for quite a range of the parameter space, but not for  $\lambda = 0$ . Figure 1 indicates that the

balanced complete bipartite network and the circle networks, can be stable for any value of  $\lambda$ .<sup>4</sup> Independent of the choice of  $\lambda$ , the circle networks are generally stable for higher cost levels, while the complete bipartite networks are stable for lower cost levels.

Of course, figure 1 does not give the complete overview of stable networks. What is not illustrated is that for most parameter settings there is a multitude of stable networks and, as Proposition 3 in the appendix shows, that there always exists at least one stable network. In particular, if the complete and the empty network are not stable, the star network has to be stable. In the next section, we identify for small network size, other networks that are stable for various settings of the parameters.

### 3.2 Enumeration

A computer can enumerate all possible networks for small  $n$  (say  $n \leq 10$ ) and check whether they are stable. To provide an overview of the number of stable networks that exist, we checked in which range of costs  $c$  each network is pairwise stable for a fixed  $\lambda$  and using  $M = n$ . We call networks “stable for a given  $\lambda$ ,” if there exists a positive cost range in which the network is pairwise stable. By excluding those networks that are only stable for an infinitely small cost range, we do not expect to lose reasonable candidates for the emerging networks, because the networks would lose their stability due to the smallest perturbations in the cost  $c$ .

Table 1 shows the number of all different stable networks found with this procedure.<sup>5</sup>

Table 1  
Number of stable networks in dependence of size.

Network size $n$	5	6	7	8
Non-isomorphic networks	34	156	1,044	12,346
Fraction of stable networks for $\lambda = 0.5$	26%	13%	4.3%	0.95%
For Closeness	6	12	21	45
For $\lambda = 0.5$	9	20	45	117
For Betweenness	4	9	18	37

<sup>4</sup> Results might look slightly different for very small network size and networks with an odd number of actors.

<sup>5</sup> We applied the enumeration to all networks of size  $n = 3, \dots, 8$  and for eleven values of  $\lambda = 0, 0.1, 0.2, \dots, 0.9, 1$ . With this procedure, we find all stable networks, except those that are stable for some not used  $\lambda$  (say  $\lambda = 0.724$ ) and not stable for all the used  $\lambda$ 's. This number of networks is likely to be small because most of the stable networks we find are stable for multiple values of  $\lambda$ .

While the number of possible networks explodes with the size  $n$ , the number of stable networks increases much more gradually. So our notion of stability – despite being a minimal requirement – can already exclude many networks from being part of a prediction. Beyond the sheer number, the enumeration can identify the stable networks. For instance, figure 2 shows the nine networks with five actors that are stable for  $\lambda = 0.5$  as well as the six networks that are stable for  $\lambda = 0$  and the four networks that are stable for  $\lambda = 1$ .

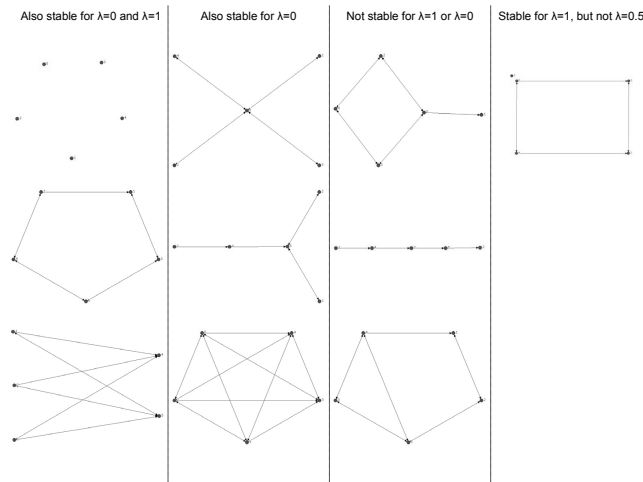


Fig. 2. All stable networks for  $\lambda = 0.5$  with indication of stability for  $\lambda = 0$  and  $\lambda = 1$ .

The enumeration provides a first indication why it is important not to concentrate on betweenness or closeness alone. The set of networks that is stable for an intermediate value of  $\lambda$  is larger than for the extreme values of  $\lambda$ . In addition, there are stable networks that are only stable for intermediate values of  $\lambda$ . This illustrates that studying stability based on one utility aspect in isolation can provide limited predictive power when multiple utility arguments are relevant simultaneously. We come back to this point below.

While the enumeration provides a full picture of the candidates for emerging networks, it does not reveal which networks are most likely the endpoint of the dynamic process. We use simulation to elaborate on the expected structural features of emerging networks for different parameter values.

### 3.3 Simulation

The simulation method follows a sequence of actions towards some stable network (cf. Doreian, 2006, Willer, 2007). To run a simulation we fix the parameter  $\lambda$  and costs  $c$ . Then, the simulation takes the following steps:

- (1) Start with some network.
- (2) Pick a pair of actors  $\{i, j\}$  at random (every pair with equal probability).
- (3) If  $ij$  does not exist, form the link  $ij$  if both  $i$  and  $j$  improve their utility (at least one strictly); if the link  $ij$  exists, sever the link if either  $i$  or  $j$  improves strictly by severing it; keep the current status of the link in all other cases.
- (4) Go back to step 1 and repeat the steps for the actual new situation until no pair of actors wants to change anymore.

We ran the simulation for 20 combinations of five levels of  $\lambda$  and four cost levels.<sup>6</sup> As starting networks we took all non-isomorphic networks for network size 3 to 8 and a sample stratified by density of around 2,500 networks for network sizes 14 and 20. Each starting network was used at least twice for each parameter combination.

A more extensive description of how such a simulation works can be found in Hummon (2000) or Buskens and Van de Rijt (2008). Its purpose is two-fold. First, for small network size, where we know all stable networks from enumeration, we use it to attach probabilities of emergence. For instance, for  $n = 8$  there are around 10 to 20 stable networks for each parameter combination. Figure 3 shows the most frequently emerging networks and their probabilities of occurrence, when starting with every non-isomorphic network three times.

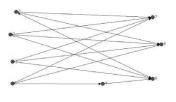





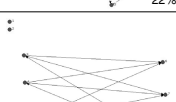
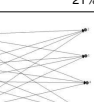
	low costs	medium costs	high costs (epsilon costs)
Closeness	 20%	 24%	 12%
0.5	 22%	 38%	 21%
Betweenness	 62%	empty network 99.9%	 49% (not high costs, but epsilon costs!)

Fig. 3. Most frequently emerging networks for  $n = 8$ .

Several observations can be made. The most frequently emerging network becomes sparser with increasing costs. For betweenness incentives, the empty network emerges with high probability already at medium cost levels. At

<sup>6</sup> The details can be requested from the authors. Costs were chosen according to analytical considerations and fit best for  $n = 8$ . For any simulation, we set  $M = n$ .

costs smaller than any positive benefit (“epsilon costs”) the complete balanced bipartite network emerges, while the complete network is unique for  $\lambda = 0$ . There are isolates when betweenness is important, but larger components do not remain disconnected, because connecting larger components increases also betweenness for the connecting actors.

The second purpose of the simulation is to run computational experiments. Starting with the same network structures, but using different utility parameters provides important insight, how changes in the utility of actors affect the emerging network structure. In the following we employ all three methods presented here (analytic analysis, enumeration, and simulation) to answer specific questions about the consequences of closeness and betweenness incentives on the network structure.

## 4 Closeness versus Betweenness Incentives

This section first contrasts the closeness dynamics with betweenness dynamics and then turns to the combination of both.

### 4.1 *Dynamics of Closeness*

For pure closeness incentives ( $\lambda = 0$ ), actors compare the costs of linking to benefits that are derived from short paths. This can be expected to be very similar to the connections model, introduced in Jackson and Wolinsky (1996).<sup>7</sup> The most prominent result of the connections model is the stability of the star network. We have shown in section 3.1 that the star is stable for pure closeness incentives in a certain cost range. Since there are also other stable networks (see also Hummon, 2000), we use the simulation to check whether the star is a likely outcome of closeness dynamics. Table 2 shows the frequency that the star emerged when starting with every non-isomorphic network of size 8 three times. This is rarely the case. Therefore, we also want to check for “star-like” networks.

The star belongs to the family of trees. These networks are characterized by being connected, but with the minimal number of links ( $n - 1$ ). To have a reference point, only 253 of the 12,346 non-isomorphic networks are trees

---

<sup>7</sup> Consistently, Borgatti and Everett (2006) list the benefits of the connections model among the “closeness-like” centrality indices. In Buechel (2008) it is shown that, indeed, the analytical results of the (symmetric) connections model correspond one-to-one to the pure closeness model. Moreover, an enumeration yielded that the set of stable networks for both models are almost coinciding.

(that is 0.2%). The enumeration for pure closeness reveals that in the set of stable networks in fact 42% are trees, while for larger weights on betweenness ( $\lambda = 0.1, 0.2, \dots, 1$ ), this is not above 22%. This indicates that trees are the dominant outcome for the dynamics of closeness although the star, their prominent representative, is not. The second row of table 2 confirms this conjecture for high costs. For lower costs, long branches are closed to become cycles. When costs are even lower, cycles are cross-connected leading to even shorter distances (see figure 3).

Table 2

Fraction of trees emerging for closeness incentives (simulation for  $n=8$ ).

	low cost	medium cost	high cost
Stable Networks	12	10	20
Trees emerging	1.0%	11.4%	90.7%
Star emerging	1.0%	0.6%	0.1%

#### 4.2 Dynamics of Betweenness

For pure betweenness ( $\lambda = 0$ ), every actor is striving for brokerage opportunities. A similar model was studied by Buskens and Van de Rijt (2008) operationalizing Burt’s idea of structural holes. In their model, complete bipartite networks are the most likely outcome of network dynamics.

In our model, it can be shown that any complete bipartite network can only be stable for the low cost part of the parameter space (as depicted in figure 1; see also proposition 4 in the appendix). Nonetheless, as the enumeration shows there are not many stable networks above this range: only 9 out of 37 stable networks for  $n = 8$ . Of the other 28 stable networks, many resemble complete bipartite networks. Some of them do not belong to this class in a strict sense, for example, a network with two isolates and a (4:2)-complete-bipartite component.

We ran a simulation for three cost levels where complete bipartite networks are stable. As table 3 shows, the family of complete bipartite networks are the dominant architecture for pure betweenness. For epsilon costs rather the connected ones emerge; for slightly higher costs rather the balanced ones. The balanced complete bipartite network, belonging to both subclasses, is the most frequently emerging network. It is notable that the empty network emerges in 99.9 percent of the simulation runs for medium and high costs, while also the circle is stable.

Summarizing, we find that the results of our model correspond with the literature on similar models. The results suggest as a rule of thumb that incentives for short paths (here closeness) lead to tree networks and incentives



Table 3

Fraction of complete bipartite networks (CB) emerging (simulation for  $n=8$ ).

	epsilon costs	very low costs	low cost
Stable Networks	19	9	4
All CBs with or without isolates	40.4%	78.3%	61.1%
CBs (2:6, 3:5, 4:4) without isolates	29.0%	38.7%	0.9%
Balanced CBs (4:4, 3:3, 2:2) with or without isolates	13.4%	37.6%	61.1%
Balanced CB (4:4) without isolates	12.5%	25.4%	0.9%

for brokerage (here betweenness) lead to complete bipartite networks at least for specific levels of tie costs.

Having characterized the emerging networks for pure closeness incentives and for pure betweenness incentives, the next question is how those results carry over to a scenario with “mixed” incentives.

### 4.3 Interaction of Closeness and Betweenness

Running simulations reveals that trees and complete bipartite networks might also emerge quite frequently for mixed incentives. For example, in a simulation with  $n = 8$ ,  $\lambda = 0.5$ , and very low costs, complete bipartite networks emerge in 37.8% of the runs; at high costs, trees emerge in 78% of the runs. This, however, is not the complete story.

By enumeration we can compare all stable networks for different incentives. Figure 4 depicts the number of stable networks for different  $\lambda$ 's. The networks are colored by the range of  $\lambda$  for which they are stable. Strikingly, there are more stable networks for each level of mixed incentives than for pure incentives. All networks that are stable for closeness incentives ( $\lambda = 0$ ) are also stable for some other  $\lambda$ . Eight of them can be stable for any  $\lambda$ ; fifteen are stable for any, but pure betweenness. For pure betweenness ( $\lambda = 1$ ), there are 37 stable networks. Fifteen of them would never occur for any other  $\lambda$  (we used). This is remarkable, as only three networks of the other categories are found stable for only one  $\lambda$ . The other stable networks for betweenness are typically also stable for any other  $\lambda > 0$ . Thus, there is strong indication that the stable networks across certain  $\lambda$ 's do not differ heavily, except for the case of pure incentives. Since most of the stable networks are neither stable for pure closeness nor for pure betweenness but for mixed incentives, many candidates of emerging networks are not covered by pure incentives. Measuring certain properties of the set of stable networks confirms that pure incentives are special cases. For example, most of the stable networks for mixed incentives are connected. For pure incentives (i.e. betweenness) this is not necessarily true: Among the stable networks for  $\lambda = 1$ , eighteen are unconnected; for  $\lambda = 0.9$ , this reduces to three (enumeration for  $n = 8$ ).

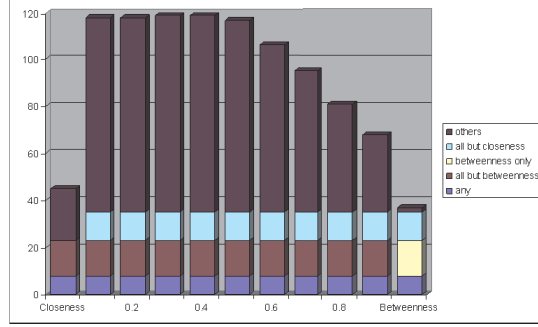


Fig. 4. Number of stable networks by  $\lambda$  (enumeration  $n = 8$ ).

So we observe that, although the weighting of benefits in our model is smooth (the benefits are a linear combination of closeness and betweenness), the results exhibit jumps. Thus, introducing a bit of betweenness (closeness) incentives into a pure closeness (betweenness) model heavily affects the results, changing and increasing the number of the candidates for emerging networks. To understand why this happens, we analyze the interaction of closeness and betweenness incentives focusing on one structural feature: the integration of isolates.

**Example 1** Consider a network  $g$  with an isolated actor  $i$  and an actor  $j$  who is already part of a larger group. Then, when closeness only matters ( $\lambda = 0$ ), actor  $i$  has a strong interest in the link  $ij$ , as this link is his first connection to the network (without  $ij$ ,  $CLOSE_i(g) = 0$ ). Actor  $j$ 's interest is restricted: creating  $ij$  means being directly connected to  $i$ , but does not have an impact on any other distance. So for high enough linking costs,  $i$  is willing to link with  $j$ , but  $j$  rejects this offer. When betweenness only matters ( $\lambda = 1$ ), actor  $j$  has a high interest in the link  $ij$ , because it provides a substantial amount of betweenness. On the other hand,  $i$  is not interested in this link as his betweenness is zero with or without  $ij$ . So the link will not be formed. Finally, when both incentives matter ( $\lambda \in [\epsilon, 1 - \epsilon]$ ), the link can be formed because both actors do have a rather high interest in this link, but for different reasons:  $i$  wants to have access to the community (closeness incentives);  $j$  enjoys mediating  $i$  with all his connections (betweenness incentives).

The example explains why for pure incentives many networks fail to be stable. Especially for pure betweenness, many networks consist of actors that do not have any incentive to keep a link. Introducing a bit of closeness benefits can justify keeping these relationships. This example finds its formal expression in lemma 1 in the appendix. Networks that consist of pendants (actors with degree one) are “more” stable for mixed incentives than for pure incentives. This is illustrated in figure 1 where the upper bound for the stability of the star network marks the upper bound for the emergence of any network with pendants (as shown in proposition 5). This area has a “wedge”-like

shape, being highest for a combination of closeness and betweenness incentives. Moreover, the example shows how different incentives can be at work, although all actors do have the same preferences.

## 5 Social Consequences

Individual goal-oriented behavior might have (possibly negative) consequences on outcomes at the collective level. We analyze two of such aspects. First, we analyze how centrality incentives determine patterns of clustering. Second, we assess the extent to which actors actually are able to establish closeness and betweenness given that they all strive for these forms of centrality.

### 5.1 Lack of Clustering

The work of Watts and Strogatz (1998) has drawn attention to clustering, a structural feature of many social networks. Small groups of actors are heavily linked among themselves such that “a friend of a friend is very likely to be also my friend”. Formally, the clustering coefficient of an actor  $i$  is defined as the number of links among his neighbors  $N_i(g)$  as a fraction of all possible links among them,  $Clust_i(g) := \frac{2\zeta_i(g)}{(l_i(g)-1)l_i(g)}$ , with  $\zeta_i(g) := \{jk \in g | j, k \in N_i(g)\}$  (see Watts and Strogatz, 1998 or Watts, 1999).

The question arises whether such patterns persist when actors start to optimize their centrality. It is the argument by Granovetter (1973) and Burt (1992) that it is not those complete triads that are the source of the many network benefits, but the open triads leading into different areas of the network. Betweenness incorporates that idea by measuring the brokerage of a given node; but also closeness favors actors at distance two rather than one if ties are sufficiently costly. So we would expect that actors optimizing their centrality will replace links in complete triads by links that bridge higher distances.

Proposition 6 in the appendix excludes the emergence of networks with high clustering among some actors. The first part shows that networks with cliques (fully connected groups of actors) cannot be stable for pure closeness. The second part shows that if an agent has a clustering coefficient of one, then the network can only be stable for very low costs and sufficient weight on closeness incentives. The reason is that full clustering implies that an agent has zero betweenness. As a consequence there is only a small cost range between the costs where the complete network is uniquely stable and the costs where no full clustering can occur. Although those results provide strong evidence against full clustering, centrality incentives might still be consistent with some

clustering. This is also ruled out by the enumeration and simulation.

There are different ways to combine the individual clustering coefficients into one. We chose a common transitivity index (see, e.g., Frank and Harary, 1982), namely, the proportion of complete triads among the triads with two or three links.<sup>8</sup> We have already shown in section 4 that many stable networks for extreme weights are complete bipartite networks or trees. These networks do not have any complete triad and, therefore, their (aggregate) clustering coefficient equals 0. Tables 4 and 5 show the results for all stable networks of size 8 and the emerging networks in a simulation with size 14.

Table 4

Stable Networks without any Complete Triad ( $n=8$ ).

	fraction
All non-isomorphic networks	3.3%
Closeness	93%
$\lambda = 0.5$	60%
Betweenness	54%

Table 5

Clustering Coefficient of Emerging Networks (simulation for  $n=14$ ).

	very low costs	low costs	medium cost
Closeness	0.19%	0.19%	0
$\lambda = 0.5$	4.86%	0.02%	0
Betweenness	0.49%	0	0

Table 4 shows that in most of the stable networks, there are no complete triads and, therefore, the clustering coefficient is zero, while in most existing networks there are complete triads. Table 5 shows the simulation result for  $n = 14$  where the starting networks have on average a clustering coefficient (transitivity index) of 66 percent. We observe that these networks with considerable clustering are transformed into networks with (almost) zero clustering by the dynamics of centrality incentives.

Summarizing we find that both closeness and betweenness incentives, destroy clustering patterns of a given network. This is consistent with the results of Holm and Ghoshal (2008) who address similar questions within a different framework. Therefore, it is likely that in most empirical settings additional phenomena are at work than that the dynamics are driven by centrality alone. First, the opportunity to meet somebody increases the likelihood to become a friend of a friend. Second, there are differences among actors, not captured by network statistics only. Geographically close actors might face lower costs maintaining relationships or other characteristics can lead to attractiveness (like homophily). Third, even if networks are determined by optimizing actors who are all the same, our benefit function does not incorporate the utility actors derive from having closed triads (i.e., strong ties, see Krackhardt, 1992). Also Burger and Buskens (2006) argue that there are cooperative contexts

<sup>8</sup> This is essentially the same measure as used as a clustering coefficient by Newman et al. (2002). An alternative is averaging the clustering coefficient of all individuals. However, by taking this average the weight of an actor with few ties is the same as the weight of an actor with many ties.

where closure is more valuable and there are competitive contexts where open networks are more beneficial.

Although it might not be that surprising result that actors striving for centrality destroy the clustering patterns of a network, the next section shows that it is also not so evident that if actors strive for centrality they actually obtain high levels of centrality.

## 5.2 Inefficiency of Stable Networks

In the economics literature on network formation considerable interest is given to the “tension” between stability and efficiency (in the sense of maximal sum of utility; see Jackson and Wolinsky, 1996 and Jackson, 2004). We also find such a tension in our model, as the individual optimizing behavior does typically not lead to a collective optimum.<sup>9</sup>

### 5.2.1 A Paradoxical Observation

It turns out that actors do not reach optimal levels of closeness and betweenness in our model when striving for closeness and betweenness, respectively. Figures 5 and 6 illustrate this by depicting the distribution of the average closeness and betweenness for all stable networks of size 8 for the eleven values of  $\lambda$  used before.

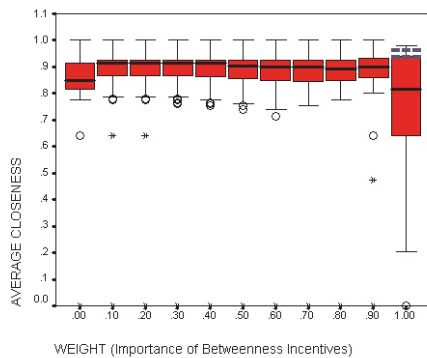


Fig. 5. Average closeness of all stable networks (enum.  $n = 8$ ).

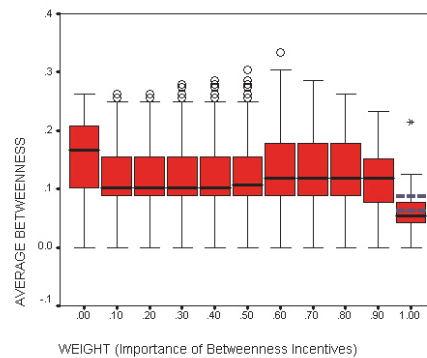


Fig. 6. Average betweenness of all stable networks (enum.  $n = 8$ ).

Paradoxically, average closeness is not highest for pure closeness incentives and average betweenness is not highest for pure betweenness incentives. Rather the

<sup>9</sup> The assessment of the tension in our model requires an elaboration of the connectedness, density, and distances of the emerging networks. This can be obtained from the authors. Within these pages we only present the main intuition and give a proof for the case  $\lambda = 1$  (see appendix proposition 7).

stable networks for closeness incentives exhibit higher average betweenness and the stable networks for betweenness incentives exhibit higher average closeness, if they are connected (the dashed lines depict the results for the subset of connected stable networks).<sup>10</sup> Moreover, by randomly picking a network from the set of all networks one finds higher average closeness (betweenness) than for the stable networks for closeness (betweenness) incentives.

To shed some light on this puzzle we derive the determinants of average closeness and average betweenness.

### 5.2.2 The Role of Distances

When computing the average closeness (of all actors) in a network, one has to consider for each actor the distance to any other actor. Denote by  $SD := \sum_{j < k} d_{jk}$  the sum of all distances in a network. Then average closeness can be written as  $AV'CLOSE = \frac{M}{M-1} - \frac{2SD}{(n-1)(M-1)n}$ , a linearly decreasing function in  $SD$  - going from 1 for the complete network to 0 for the empty network. So average closeness of a network is fully determined by the sum of distances.

Although this is not so easy to see, average betweenness of the actors is also a function of the sum of distances. In the appendix proposition 8 shows that  $AV'BETWEEN = \alpha \sum_{j < k: \text{connected}} (d_{jk} - 1)$ , a linearly increasing function of the distances among all connected agents (where  $\alpha > 0$  is a constant). Average betweenness is zero for the empty and the complete network, and is maximal for the network with maximal distances - the line network. The maximal average betweenness is  $\frac{1}{3}$ , not 1, as there is no network with every pair of actors at the maximum distance. Arguably, the brokerage benefits that a pair offers are rather constant than increasing in their distances such that being on a long shortest path, is not as beneficial as being on a short one (see, e.g., Goyal and Vega-Redondo, 2008). But this is not true for the standard definition of betweenness (see, e.g., Wasserman and Faust, 1994) used here.

This implies for connected networks that both, average closeness and average betweenness are fully determined by the sum of all distances. Since the former is linearly decreasing and the latter is linearly increasing in sum of distances, this implies that for connected networks average closeness and average betweenness are perfectly negatively correlated.<sup>11</sup> So networks with high average closeness are those that contain short distances and networks with high average

<sup>10</sup> Not to misinterpret: this observation does not imply that if an actor wants to maximize his betweenness, he should strive for closeness or the other way around.

<sup>11</sup> In this paper we work with closeness as the reverse distance, not the inverse distance as usual. Using Freeman's definition of closeness, the negative correlation between average (Freeman) closeness and average betweenness is also almost perfect, namely, -0.978.

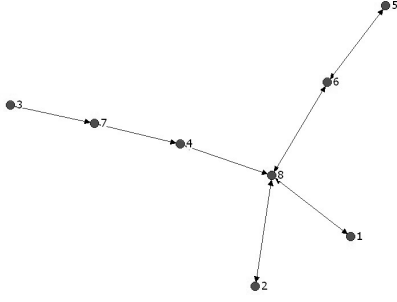


Fig. 7. Very frequently emerging network for closeness incentives (simulation  $n = 8$ ).

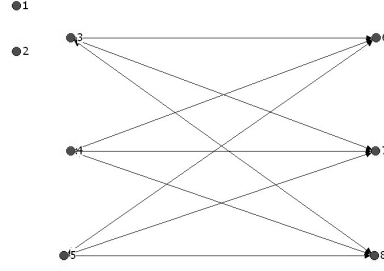


Fig. 8. Very frequently emerging network for betweenness incentives (enumeration  $n = 8$ ).

betweenness have long distances but are connected. This insight transforms the efficiency puzzle into the puzzle why distances are relatively large under pure closeness incentives and why they are relatively long under pure betweenness incentives. Indeed, when searching for stable networks that have a smaller average distance than the average of all non-isomorphic networks (not counting  $M$  for unconnected actors), then we find 3 out of 45 for pure closeness and 26 out of 37 for pure betweenness (enumeration for  $n = 8$ ).

### 5.2.3 The External Effects

Some part of the puzzle is due to cost effects. As for low costs, networks are denser and have shorter distances than for high costs, it happens that low costs lead to networks with high average closeness, while high costs lead to high average betweenness under the condition that the resulting network is connected. However, when controlling for costs some inefficiency remains. This can be illustrated in two examples, one for pure betweenness and one for pure closeness.

**Example 2** *Figure 7 depicts a network that is frequently emerging for pure closeness (found in the simulation). This network is stable if  $lb \approx 0.12245 < c \leq ub \approx 0.14268$ . The additional costs are maximally  $2ub$  (once for each holder). Socially, it is beneficial to take these costs if the total closeness benefits increase sufficiently. That is  $2ub > \Delta CLOSE \Leftrightarrow \Delta SD > 7$ . In this network there are multiple possibilities to shorten the sum of distances by more than 7, e.g., a link between actors “5” and “7” would reduce the sum of distances by 9. However, the network is stable. As this problem occurs even at the upper bound of the costs, this network is “underconnected” whenever it emerges, in the sense that the addition of links would socially be an improvement.*

**Example 3** *Proposition 8 has shown that the networks with the highest betweenness are sparse, exhibiting long distances. For instance, the network in figure 7 exhibits very high average betweenness. However, it is not stable for*

*pure betweenness. By the creation of a link (e.g., between actors “5” and “7”) the focal actors can substantially increase their individual betweenness, while this would do collective harm, in the sense that it lowers average betweenness. And the evolution will not be finished after one change: more links will be added, and also links will be cut (by actors that have no betweenness flow from their ties). As the endpoint typically a network like in figure 8 emerges. In this network there are no substantial brokerage benefits left: the average betweenness is lower than in an arbitrary network, and much lower than in the network of figure 8.*

The two examples above illustrate two different types of inefficiency. In the first example the network is sparse consisting of “too long” distances (figure 7); while in the second example the network consists of a component that is “too dense” (figure 8) with insufficient distances. This can be shown to be a genuine difference between closeness incentives and betweenness incentives: Consider two actors  $i$  and  $j$  (at some distance  $1 < d < M$ ) who decide to form a tie. What does that agreement mean for the other actors? First of all, their costs have not changed. Secondly, they might improve their closeness, because distances might reduce, but cannot be increased. Finally, by proposition 8 average betweenness is certainly lowered. While the betweenness of  $i$  and  $j$  can stay constant or increase by the formation of the tie, there must be other actors who lose betweenness benefits. This is plausible, because the link  $ij$  first of all takes away the brokering benefits for all actors that were on their geodesics before. Moreover  $i, j$  can now be on shortest paths where others were before.

The key point is that the costs of a new link are fully internalized by the focal actors, while the benefits have spill-overs: positive externalities on closeness and mainly negative externalities on betweenness. The actors of our model, however, do not consider these effects when changing the network structure. As a consequence emerging networks for closeness incentives are often “too” sparse as in example 2, while emerging networks for betweenness incentives are “too dense” in respect to average betweenness (proposition 7 in the appendix shows that this problem always occurs for pure betweenness).

## 6 Concluding Remarks

The innovations of this paper are three-fold: (a) Although both betweenness and closeness centrality are cornerstones of social network analysis, it has not been explicitly studied which networks will emerge if agents follow incentives for these two positional advantages. We can relate the dynamics of closeness to existing models where actors strive for short paths and the dynamics of betweenness to models of brokerage opportunities. By also including costs for



the number of ties, we include degree centrality, the third centrality measure from the classic article by Freeman in 1978. (b) There is hardly any theoretical work that studies the interplay between different types of incentives to predict network formation processes. When combining incentives for closeness and betweenness, we find results that are not straightforward extension of considering them separately. We explain this phenomenon based on the observation that two agents, despite similar preferences, can have quite different motivations of action, based on asymmetry in their network positions. (c) We draw attention to the social consequences of individual optimizing behavior in social networks. First, there is support for the claim that centrality driven decisions eliminate structural patterns that are typical for friendship networks (i.e., clustering). More surprisingly, it is not likely that actors striving for closeness or betweenness benefits reach a social optimum as networks with, respectively, relatively low average closeness and low average betweenness emerge. The discrepancy between individual incentive and collective outcome is caused by two different types of social dilemmas.

The first dilemma has the flavor of a  $n$ -person prisoners dilemma (also labeled as public goods problem) and occurs when actors strive for short paths, operationalized as closeness centrality. The creation of a tie exhibits positive externalities as not only the distances of the active actors changes, but also some other agents benefit from this action. When facing costs of link maintenance, actors refuse to be the ones who “produce this public good”. Consequently, the emerging networks are often too sparse compared with what would have been efficient. More specifically, we find especially for higher costs of establishing relations that many tree networks emerge that typically have rather large distances.

The second dilemma has the flavor of the “tragedy of the commons” and occurs when actors strive for betweenness centrality. Two actors creating a tie can often improve their betweenness position substantially. What they do not consider is the effects this action has on other actors: this is typically a negative effect (in the sense of reducing average utility). Consequently, the emerging are too dense compared with what would have been efficient. Especially with low tie costs, complete bipartite networks emerge in which betweenness is rather low for all actors.

By also considering mixed closeness and betweenness incentives, this paper extends earlier research in this field. Under mixed incentives, many networks can emerge that do not emerge under pure closeness and betweenness incentives. Moreover, networks in these mixed settings are rather efficient, since the two problems of inefficiency balance one another.

Clearly, we also neglected various other aspects of network dynamics that might be incorporated in future research. For example, we restrict the benefits

of mating to “structural goals” (see Doreian, 2006). By ignoring actor specific and dyadic specific explanatory variables, we exclude effects such as homophily. Further effects are excluded by studying only non-directed networks, i.e. effects of reciprocity, and by considering a competitive setting as opposed to a cooperative setting, i.e. effects of closure (see Burger and Buskens, 2006 for a comparison of these two settings). These might be reasons for why we find low clustering in emerging network in contrast with rather some empirical observations of networks. Finally, we work with high information assumptions: actors are fully informed about the network and able to optimize their utility given this information. In further research it is worthwhile to study whether less-informed actors following simple behavioral rules induce different dynamics.

## A APPENDIX

In the appendix we provide formal statements and proofs for the propositions that we discussed throughout the main text. The results are based on the centrality model we explained in section 2.3. Using the definition of closeness and betweenness (equations 1 and 2), actor utility was defined by:

$$u_i(g) = (1 - \lambda)CLOSE_i(g) + \lambda BETWEEN_i(g) - cDEGREE_i(g). \quad (A.1)$$

Analytical results on stability typically need the maximal incentive of any actor to sever a link and the maximal incentive of any two actors to add a link and compare them to linking costs  $c$ . Because benefits are based only on closeness and betweenness, the crucial aspects for a focal actor  $i$  are the change in **distances**  $\sum_{j \in N} d_{ij}(g)$  ( $\hat{=}$  non-normalized closeness) and the change in the number of shortest paths he is on  $\sum_{j < k (j \neq i, k \neq i)} \frac{\tau_{jk}^i(g)}{\tau_{jk}(g)}$  ( $\hat{=}$  non-normalized betweenness), what we call his “**brokerage**”. It is sometimes left to the reader to plug in these changes into the utility function above. Specifically, *if a new link for some actor  $i$  in some network  $g$  means a decrease in distances of  $X$  and an increase in brokerage of  $Y$ , then he is willing to form the link only if*  $c \leq \frac{(1-\lambda)[X]}{(M-1)(n-1)} + \frac{\lambda 2[Y]}{(n-1)(n-2)}$ .

Although deriving the changes in distances and brokerage for a given situation might be tedious, it is mostly a straightforward task. The results are ordered according to their appearance in the main text.

### A.1 Proofs of section 3

**Proposition 1** *In the centrality model the following holds:*

- (1) The complete network  $g^N$  is stable if and only if  $c \leq \frac{1-\lambda}{(n-1)(M-1)}$ .
- (2) The empty network  $g^\emptyset$  is stable if and only if  $c \geq \frac{1-\lambda}{n-1}$ .
- (3) A star shaped network  $g^*$  is stable if and only if  $\frac{1-\lambda}{(n-1)(M-1)} \leq c \leq \min\left\{\frac{1+\lambda}{n-1}, \frac{(1-\lambda)[M(n-1)-2n+3]}{(n-1)(M-1)}\right\}$ .
- (4) Let  $n$  be a multiple of 4. Then the circle network  $g^\circ$  is stable if and only if  $\frac{(1-\lambda)[\frac{1}{8}n^2 - \frac{1}{2}n + 1]}{(M-1)(n-1)} + \frac{2\lambda[\frac{1}{8}n^2 - \frac{3}{4}n + 1]}{(n-1)(n-2)} \leq c \leq \frac{(1-\lambda)[\frac{1}{4}n^2 - \frac{1}{2}n]}{(M-1)(n-1)} + \frac{2\lambda[\frac{1}{8}n^2 - \frac{1}{2}n + \frac{1}{2}]}{(n-1)(n-2)}$ .
- (5) A complete bipartite network  $g^{l:r}$  with  $2 \leq r \leq l$  (where  $l$  and  $r$  are the sizes of the two groups) is pairwise stable if and only if  $\frac{1-\lambda}{(n-1)(M-1)} \leq c \leq \frac{2(1-\lambda)}{(n-1)(M-1)} + \frac{2\lambda[\frac{r-1}{l}]}{(n-1)(n-2)}$ .

**Proof.** The results of proposition 1 present lower and/or upper bounds of costs where a network is claimed to be stable. For conciseness, we denote with  $lb(g)$  the claimed lower bound of a network  $g$  and analogously the claimed upper bound with  $ub(g)$ .

- (1) The complete network  $g^N$  can only be altered by deletion of a link. Any actor deleting any link increases his distances by 1 and does not change his brokerage. Therefore, no actor will sever a link for  $c \leq ub(g^N)$  and any actor wants to sever a tie for higher costs.
- (2) The empty network  $g^\emptyset$  can only be altered by the addition of links. Any actor adding any link decreases his distances by  $M-1$ , while his brokerage remains zero. Thus, no actor will do that for  $c \geq lb(g^\emptyset)$  and any pair of actors is willing to add a tie for  $c < lb(g^\emptyset)$ .
- (3) Only peripheral actors can add links. Any actor adding a link reduces his distances by 1 and does not change his brokerage. This leads to the  $lb(g^*)$ . The central actor severing a link increases his distances by  $M-1$  and decreases his brokerage by  $n-2$ . A peripheral actor cutting a link increases his distances by  $M-1 + (n-2)(M-2)$  and does not change his brokerage. Plugging into the utility function yields that no actor wants to sever a link for  $c \leq \min\{ub1(g^*); ub2(g^*)\}$ , while some actor is willing to sever a tie for higher costs.
- (4) Any actor severing any link increases his distances from the circle to the line network. For  $n$  even this is a change in distances of  $\frac{1}{4}n^2 - \frac{1}{2}n$  and a change in brokerage from  $\frac{1}{8}n^2 - \frac{1}{2}n + \frac{1}{2}$  to zero, yielding the upper bound. Two actors forming a link benefit the further away they are. For  $n$  a multiple of four, two actors on opposite sides (with two shortest paths) can form a link building a network with two cycles of odd length. Their change in distances can be derived as  $\frac{1}{8}n^2 - \frac{1}{2}n + 1$ , while their brokerage changes by  $\frac{1}{8}n^2 - \frac{3}{4}n + 1$ . In the same way slightly different inequalities can be derived for other network sizes.
- (5) In complete bipartite networks, additional links are only possible within a group. Since everybody is already indirectly linked, any actor adding a link reduces his distances by 1 without changing his brokerage. This

yields the  $lb(g^{l:r})$ .

Since both groups consist of at least two actors, cutting one link only affects the distance between the focal actors. Their distance changes by 2. The brokerage for an actor in the group of size  $l$  changes by  $\frac{1}{l}(r-1)$ , because he is on one of  $l$  shortest paths between any pair of actors in the other group. Because  $l \geq r$ , the actors in the  $l$ -group benefit less from their links. They are indifferent about cutting if  $c = \frac{2(1-\lambda)}{(n-1)(M-1)} + \frac{2\lambda \lceil \frac{r-1}{l} \rceil}{(n-1)(n-2)} = ub(g^{l:r})$ . Therefore, for  $c < lb(g^{l:r})$ , two actors of the same group form a link; for  $c > ub(g^{l:r})$  an actor of the larger group (of size  $l$ ) will sever a link. And, no actor can improve by changing a link for  $lb(g^{l:r}) \leq c \leq ub(g^{l:r})$ .

Plugging in  $l = r = \frac{n}{2}$  yields that a balanced complete bipartite network  $g^{\frac{n}{2}:\frac{n}{2}}$  (for even  $n$ ) is pairwise stable if and only if  $\frac{1-\lambda}{(n-1)(M-1)} \leq c \leq \frac{2(1-\lambda)}{(n-1)(M-1)} + \frac{2\lambda \lfloor 1 - \frac{2}{n} \rfloor}{(n-1)(n-2)}$ .  $\square$

**Proposition 2** *In the centrality model the complete network  $g^N$  is uniquely stable if and only if  $c < \frac{1-\lambda}{(n-1)(M-1)}$ .*

**Proof.** The proof is analogue to proposition (6i) of Buechel (2008). The complete network  $g^N$  is stable because  $c \leq \frac{1-\lambda}{(n-1)(M-1)}$  (see proposition 1). For any network  $g \in \{G \setminus g^N\}$ ,  $\exists(i, j) : d_{ij}(g) > 1$ . By connecting the distances of  $i$  and  $j$  decrease at least by 1. So for  $c < \frac{1-\lambda}{(n-1)(M-1)}$  they want to connect and, therefore, the network is unstable. The complete network is not uniquely stable for other values of  $c$  because if  $c \geq \frac{1-\lambda}{(n-1)(M-1)}$ , the star or the empty network is stable (see proposition 1).  $\square$

**Claim 1** *In the centrality model the empty network  $g^\emptyset$  typically is uniquely stable for  $c > (1-\lambda)\frac{n-1}{4(M-1)} + \lambda\frac{n^2-4n+3}{4(n-1)(n-2)}$ .*

“Typically” indicates that this statement is valid for most settings, but not for any combinations of  $n$  and  $M$ . For a justification note that this threshold for uniqueness is just  $ub(g^\circ)$  for  $n$  odd (which is slightly higher than for  $n$  even). We argue that among all networks that contain non-critical links, the circle network has the highest marginal benefit for any of those (non-critical links). This implies that for  $c > ub(g^\circ)$  in any network with cycles (cycles consist of non-critical links) at least one actor is willing to cut a link. Thus  $ub(g^\circ)$  is the maximal cost level, where any network with cycles can be stable. Then we use that a network without cycles is either empty or does contain pendants (actors of degree one). Therefore, to establish uniqueness of the empty network it remains to show that networks with pendants are not stable. For costs around  $ub(g^\circ)$  and higher this is typically, but not generally, true (see lemma 1).

**Proposition 3** *In the centrality model for any parameters  $(\lambda, c) \in [0, 1] \times \mathbb{R}_+$  there exists a stable network.*

**Proof.** This theorem follows almost directly from proposition 1. For  $\lambda = 1$ , the empty network  $g^\emptyset$  is trivially stable at any cost. For  $\lambda < 1$ ,  $g^\emptyset$  is stable if  $c \geq \frac{1-\lambda}{n-1}$ ,  $g^N$  is stable if  $c \leq \frac{1-\lambda}{(n-1)(M-1)}$  and  $g^*$  is stable if  $\frac{1-\lambda}{(n-1)(M-1)} \leq c \leq \min\{ub1(g^*) = \frac{1+\lambda}{n-1}, ub2(g^*) = (1-\lambda)[\frac{M}{M-1} - \frac{2n-3}{(M-1)(n-1)}]\}$ .

It remains to be shown that if  $g^\emptyset$  and  $g^N$  are not stable,  $g^*$  is stable. This follows directly from  $lb(g^*) = ub(g^N)$ ,  $ub1(g^*) \geq lb(g^\emptyset)$  as  $\frac{1+\lambda}{n-1} \geq \frac{1-\lambda}{n-1}$ , and  $ub2(g^*) \geq g^\emptyset$  (by definition  $n \geq 3$  and  $M \geq n-1$ , which implies  $\frac{M(n-1)-2n+3}{M-1} \geq 1$ ).  $\square$

## A.2 Proofs of section 4

**Proposition 4** *In the centrality model a complete bipartite networks  $g^{l:r}$  can only be stable if the balanced complete bipartite network  $g^{\frac{n}{2}:\frac{n}{2}}$  is stable, that is if  $\frac{1-\lambda}{(n-1)(M-1)} \leq c \leq \frac{2(1-\lambda)}{(n-1)(M-1)} + \frac{2\lambda[1-\frac{2}{n}]}{(n-1)(n-2)}$ .*

**Proof.** This result follows from proposition 1, part (5): For  $c < lb(g^{\frac{n}{2}:\frac{n}{2}}) = lb(g^{l:r})$  two actors of the same group form a link (the complete network is uniquely stable); for  $c > ub(g^{\frac{n}{2}:\frac{n}{2}}) \geq ub(g^{l:r})$  an actor of the larger group (of size  $l$ ) will sever a link.  $\square$

The following lemma 1 is helpful:

**Lemma 1** *If  $c > \min\{\frac{1+\lambda}{n-1}, \frac{(1-\lambda)[M(n-1)-2n+3]}{(n-1)(M-1)}\}$ , no network with pendants (actors of degree one) can be pairwise stable.*

**Proof.** Take any network  $g$  with a pendant  $i$  and his neighbor  $j$ . We show that the condition implies that one of the actors wants to sever this link.

- (1) Actor  $i$  does not reduce brokerage by severing this link. Removing the link increases his distances at least by  $M-1$  (when actor  $j$  is also a pendant) and at most by  $M-1 + (n-2)(M-2)$  (when actor  $j$  is directly linked to all other actors). Therefore, actor  $i$  will not keep the link if  $c > \frac{(1-\lambda)[M(n-1)-2n+3]}{(n-1)(M-1)}$ .
- (2) Similarly, for actor  $j$ : severing a link increases his distances by  $M-1$  and hence decrease his closeness by  $\frac{1}{n-1}$ . Moreover, he was on the shortest path between  $i$  and any other actor in this component. The more actors in this component, the higher the incentive to keep this link. The maximum brokerage of  $n-2$  is attained for a connected network. Therefore, actor  $j$  wants to sever the link for  $c > \frac{1-\lambda}{n-1} + \frac{\lambda 2(n-2)}{(n-1)(n-2)}$  rendering the network unstable.  $\square$

**Proposition 5** *In the centrality model a tree can only be stable if the star is stable that is if  $\frac{1-\lambda}{(n-1)(M-1)} \leq c \leq \min\{\frac{1+\lambda}{n-1}, \frac{(1-\lambda)[M(n-1)-2n+3]}{(n-1)(M-1)}\}$ .*

**Proof.** We have to justify the lower bound ( $lb$ ) and the upper bounds ( $ub1$ ,  $ub2$ ). By proposition 2 if  $c < lb$ , the complete network is uniquely stable. For the upper bounds, we use that any tree has pendants. By the lemma 1, no tree can be stable for  $c > \min\{ub1; ub2\}$ .  $\square$

### A.3 Proofs of section 5

**Proposition 6** *In the centrality model the following holds:*

- (1) For  $\lambda = 0$ , a network with a clique of size  $q (\geq 3)$  or larger cannot be stable if  $c > \frac{n}{q(M-1)(n-1)}$ .
- (2) For any network  $g$ , actor  $i$  with  $l_i(g) \geq 2$  and costs  $c > (1-\lambda) \frac{n-l_i(g)-1}{l_i(g)(M-1)(n-1)}$  it holds that, if  $i$  has full clustering ( $Clust_i(g) = 1$ ), then the network cannot be stable.

**Proof.** The two parts of the proposition are independent.

- (1) Let  $H \subset N$  be a completely linked group of size  $q$  and  $h \in G$  its links. We compare all networks that contain  $h$  in respect to the weakest link among the subgroup. More precisely, denote by  $\beta_i^{ij}(g) := CLOSE_i(g) - CLOSE_i(g \setminus ij)$ ; then we look for the

$$\max_{g: \exists h \in g} \min_{\{i,j\} \in H} \beta_i^{ij}(g). \quad (\text{A.2})$$

We claim that among the argmax is a core-periphery network with the following configuration:  $H$  is the core and each of the actors in  $H$  is a gatekeeper for a number of peripheral actors (in the sense that the peripheral actors reach other core members only via their gatekeeper). Moreover, the peripheral actors ( $N \setminus H$ ) are equally distributed among the core members (with differences of 1 if necessary).

If in such a network cutting a link  $ij \in h$  is improving for some actor  $i \in H$ , no network that contains  $h$  can be stable. For  $n$  a multiple of  $q$  the marginal benefit of each link among the core members is  $\hat{\beta}(q) = \frac{n}{q(M-1)(n-1)}$ , because their distances change by  $\frac{n}{q}$ . If  $n$  is not a multiple of  $q$ , the marginal benefit of links to at least one group member is smaller. So for marginal costs higher than this threshold, it is always worthwhile to sever a link.

- (2) Consider  $g \in G$ ,  $i \in N$  with  $l_i(g) \geq 2$  and  $Clust_i(g) = 1$ . We show that  $i$  is willing to sever a link. First, we observe that there is no betweenness incentive to keep the link. Formally, for all  $g'$  and actors  $j$  with  $l_j(g') \geq 2$ , it holds that  $Clust_j(g') = 1 \Leftrightarrow Between_i(g') = 0$  (for a proof see, e.g., Everett et al. (2004) or Gago Alvarez (2007)). It follows that  $Between_i(g) = 0$  and cannot be improved by deletion of a link (as

the clustering coefficient remains 1). So, for stability of  $g$ , there must be sufficient closeness incentive for  $i$  to keep the links.

The maximal marginal closeness  $i$  can have for all of his links is realized in a network as follows: all actors that are not linked with  $i$  are equally “distributed” among his neighbors (in the sense that each is connected to  $i$  only via one neighbor). This is the case because (a) the maximal worth an actor  $k \notin N_i(g)$  can have for  $i$  in respect to the link  $ij$  is 1 and (b) if those actors are not equally distributed, then the link that leads to the least actors is weaker than the others.<sup>12</sup> This maximal marginal benefit can be computed as  $\frac{n-l_i(g)-1}{(n-1)(M-1)l_i(g)}$ . For costs higher than this threshold,  $i$  will cut one of his links, which makes  $g$  unstable.  $\square$

**Proposition 7** *In the centrality model for  $\lambda = 1$  the following holds: any stable network that is non-empty is inefficient (in the sense that it does not maximize the average utility).*

**Proof.** By lemma 1 no stable network contains pendants for  $\lambda = 1$  (since  $c > 0$ ). Therefore, all non-empty stable networks contain non-critical links. Let  $g$  be such a network and  $ij$  be one of the non-critical links. Severing  $ij$  increases the sum of actual distances  $\sum_{j < k: \text{connected}} d_{jk}(g)$ . Proposition 8 implies that deletion of  $ij$  increases average betweenness and decreases average costs. So,  $g$  is not efficient.  $\square$

**Proposition 8** *For any network  $g$  the following holds:*

- (1) *average closeness can be written as*  

$$\frac{1}{n} \sum_{i \in N} CLOSE_i(g) = \frac{M}{(M-1)} - \frac{2 \sum_{j < k} d_{jk}(g)}{n(n-1)(M-1)}, \text{ and}$$
- (2) *average betweenness can be written as*  

$$\frac{1}{n} \sum_{i \in N} BETWEEN_i(g) = \frac{1}{n} \frac{2}{(n-1)(n-2)} \sum_{j < k: \text{connected}} (d_{jk}(g) - 1).$$

**Proof of proposition 8, part (1).** By definition of closeness (eq. 1):

$$\frac{1}{n} \sum_{i \in N} CLOSE_i(g) = \frac{1}{n} \sum_{i \in N} \left[ \frac{M}{M-1} - \frac{\sum_{j \in N} d_{ij}(g)}{(M-1)(n-1)} \right] =$$

$$\frac{1}{n} \sum_{i \in N} \frac{M}{M-1} - \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N} d_{ij}(g)}{(M-1)(n-1)}. \text{ By taking the sums we get the result:}$$

$$\frac{nM}{n(M-1)} - \frac{\sum_{i \in N} \sum_{j \in N} d_{ij}(g)}{n(n-1)(M-1)} = \frac{M}{M-1} - \frac{2 \sum_{j < k} d_{jk}(g)}{n(n-1)(M-1)}. \quad (\text{A.3})$$

To proof part (2) we first show the following lemma:

<sup>12</sup>Note that cutting a link to a neighbor ( $j$ ) cannot increase the distance to any actor by more than 1, as  $i$  is connected to other actors, which are connected to  $j$ . Hence the weakest link is the one that does not lead to a considerable number of indirect connections.

**Lemma 2**  $\forall g \in G$  and  $\forall j \neq k (\in N : \text{connected under } g)$ , it holds that

$$\sum_{i \in N \setminus \{j, k\}} \frac{\tau_{jk}^i(g)}{\tau_{jk}(g)} = d_{jk}(g) - 1. \quad (\text{A.4})$$

In words: fixing a connected pair of actors  $(j, k)$  and summing up all actors that are on one or more of their geodesics (weighted by the fraction they are on) results in counting the length of the shortest path between  $j$  and  $k$ .

**Proof of Lemma 2.** Take any  $g$  with a pair of connected actors  $j$  and  $k$ . Let  $t (\neq 0)$  be the number of geodesics  $\tau_{jk}(g) = t$  and let  $H$  be the set of actors who are on some geodesics, that is  $H = \{i \in N \setminus \{j, k\} \mid \tau_{jk}^i(g) > 0\}$ . Denote by  $h_x$  the number of distinct actors who are on  $x$  geodesics ( $h_x := \#\{i \in N \setminus \{j, k\} \mid \tau_{jk}^i(g) = x\}$ ). The definitions imply that (\*)  $|H| = h_1 + h_2 + \dots + h_t$ . Note first that if the  $t$  geodesics are independent (disjoint), then there are  $t(d_{jk}(g) - 1)$  distinct actors in  $H$ . This number is reduced by any actor that is on more than one geodesic:

$$\begin{aligned} |H| &= t(d_{jk}(g) - 1) - h_2 - 2h_3 - 3h_4 - \dots - (t-1)h_t \\ \iff t(d_{jk}(g) - 1) &= |H| + h_2 + 2h_3 + 3h_4 + \dots + (t-1)h_t \end{aligned} \quad (\text{A.5})$$

On the other hand, recall that  $\tau_{jk}^i(g) = x$  means that actor  $i$  is on  $x$  geodesics of  $j$  and  $k$ . By the definition of  $h_x$  we can write

$$\sum_{i \in N \setminus \{j, k\}} \tau_{jk}^i(g) = h_1 + 2h_2 + 3h_3 + \dots + th_t \quad (\text{A.6})$$

We show that the left-hand side (LHS) of A.5 equals the LHS of A.6 by subtracting the right-hand sides RHS (A.6)-(A.5):

$$\begin{aligned} \sum_{i \in N \setminus \{j, k\}} \tau_{jk}^i(g) - t(d_{jk}(g) - 1) &= h_1 + h_2 + \dots + h_t - |H| =^* 0 \\ \implies \frac{\sum_{i \in N \setminus \{j, k\}} \tau_{jk}^i(g)}{t} &= d_{jk}(g) - 1 \end{aligned}$$

where the “\*” part follows from the definitions.<sup>13</sup>

<sup>13</sup> The same result was also found by Gago Alvarez (2007). To check the plausibility of the lemma just let the  $t$  geodesics be fully independent. Then  $|H| = t(d_{jk}(g) - 1)$ . Each actor  $i \in H$  derives a betweenness of  $\frac{1}{t}$ . Hence  $\sum_{i \in N \setminus \{j, k\}} \frac{\tau_{jk}^i(g)}{\tau_{jk}(g)} = \sum_{i \in H} \frac{\tau_{jk}^i(g)}{\tau_{jk}(g)} = t(d_{jk}(g) - 1) * \frac{1}{t} = d_{jk}(g) - 1$ .



**Proof of proposition 8, part (2)** As the division by  $n$  cancels out it suffices to show that  $\forall g$ ,

$$\sum_{i \in N} BETWEEN_i(g) = \frac{2}{(n-1)(n-2)} \sum_{j < k: d_{jk}(g) < M} (d_{jk}(g) - 1). \quad (\text{A.7})$$

By definition of betweenness (eq. 2)

$$\sum_{i \in N} BETWEEN_i(g) = \sum_{i \in N} \left[ \frac{2}{(n-1)(n-2)} \sum_{\substack{j < k \\ (j \neq i, k \neq i)}} \frac{\tau_{jk}^i(g)}{\tau_{jk}(g)} \right].$$

By changing summation we get:

$$\frac{2}{(n-1)(n-2)} \sum_{i \in N} \left[ \sum_{\substack{j < k \\ (j \neq i, k \neq i)}} \frac{\tau_{jk}^i(g)}{\tau_{jk}(g)} \right] = \frac{2}{(n-1)(n-2)} \sum_{j < k} \left[ \sum_{i \in N \setminus \{j, k\}} \frac{\tau_{jk}^i(g)}{\tau_{jk}(g)} \right]$$

The fraction in brackets was defined to be zero if the denominator is zero. Since this is always true for unconnected pairs (i.e.  $d_{j,k}(g) = M \implies \tau_{j,k}(g) = 0$ ), only connected pairs count for the sum before the brackets. Therefore we can apply lemma 2 which yields the result:

$$\sum_{i \in N} BETWEEN_i(g) = \frac{2}{(n-1)(n-2)} \sum_{j < k: \text{connected}} (d_{jk}(g) - 1).$$

□

## References

- Bala, V., and Goyal, S., 2000. A Noncooperative Model of Network Formation. *Econometrica* 68, 1181–1229.
- Barabási, A.-L., Albert, R., 1999. Emergence of Scaling in Random Networks. *Science* 286, 509-512.
- Borgatti, S., and Everett, M., 2006. A Graph-Theoretic Perspective on Centrality. *Social Networks* 28, 466-484.
- Buechel, B., 2008: Network Formation with Closeness Incentives. In: Naimzada, A., Stefani, S., and Torriero A. (Eds.), *Networks and Dynamics*. Springer, forthcoming.
- Burger, M., and Buskens, V., 2006. *Social Context and Network Formation: Experimental Studies*. ISCORE 234, Utrecht University.
- Burt, R.S., 1992. *Structural Holes. The Social Structure of Competition*. Harvard University Press.
- Burt, R.S., 2002. Bridge Decay. *Social Networks* 24(4), 333-63.
- Buskens, V., and Van de Rijt, A., 2008. Dynamics of Networks if Everyone strives for Structural Holes. Forthcoming in *American Journal of Sociology*.
- Buskens, V., 2002. *Social Networks and Trust*. Kluwer.

- Coleman, J., 1988. Social Capital in the Creation of Human Capital. *American Journal of Sociology* 94, 95-120.
- Dekker, D., Franses, P.H., and Krackhardt, D. 2003. An Equilibrium-Correction Model for Dynamic Network Data. *Journal of Mathematical Sociology* 27, 193-215.
- Dillig, A., 2008. Gut Vernetzt. *Neon*, May 2008, 100-110.
- Doreian, P., and Stokman, F., 2003. Introduction to Special issue on Evolution of Social Networks, Part III. *Journal of Mathematical Sociology* 27, 85-87.
- Doreian, P., 2006. Actor Network Utilities and Network Evolution. *Social Networks* 28(2), 137-164.
- Erdős, P., and Renyi, A., 1960. On the Evolution of Random Graphs. *Publication of the Mathematical Institute of the Hungarian Academy of Sciences* 5, 17-61.
- Everett, M.G., Sinclair, P., and Dankelman, P., 2004. Some Centrality Results New and Old. *Journal of Mathematical Sociology* 28, 215-227.
- Frank, O., Harary, F., 1982. Cluster Inference by Using Transitivity Indices in Empirical Graphs. *Journal of the American Statistical Association* 77 (380), 835-840.
- Freeman, L., 1979. Centrality in Social Networks: Conceptual Clarification. *Social Networks* 1(3), 215-239.
- Gago Alvarez, S., 2007. The Betweenness Centrality of a Graph. Accessed 22 May 2007. Available from <https://upcommons.upc.edu/eprints/handle/2117/972>.
- Galeotti, A., Goyal, S., and Kamphorst, J., 2006. Network Formation with Heterogeneous Players. *Games and Economic Behavior* 54, 353-372.
- Goyal, S., and Joshi, S., 2006. Unequal Connections. *International Journal of Game Theory* 34, 319-349.
- Goyal, S., and Vega-Redondo, F., 2008. Structural Holes in Social Networks. *Journal of Economic Theory*, forthcoming.
- Goyal, S., 2007. *Connections: An Introduction to the Economics of Networks*. Princeton University Press.
- Granovetter, M., 1973. The Strength of Weak Ties. *American Journal of Sociology* 78, 1360-1380.
- Granovetter, M., 1995[1974]. *Getting a Job*. University of Chicago Press.
- Holm, P., and Ghoshal, G. 2008: The diplomats dilemma: Maximal power for minimal effort in social networks. Accessed 24 July 2008. Available from <http://aps.arxiv.org/abs/0805.3909>.
- Hummon, N.P., 2000. Utility and Dynamic Social Networks. *Social Networks* 22, 221-249.
- Jackson, M., and Wolinsky, A., 1996. A Strategic Model of Social and Economic Networks. *Journal of Economic Theory* 71, 44-74.
- Jackson, M., 2004. A Survey of Models of Network Formation: Stability and Efficiency. In: Demange, G., and Wooders, M. (Eds.), *Group Formation in Economics; Networks, Clubs and Coalitions*. Cambridge University Press, chapter 1.

- Kamphorst, J., 2007. Network Formation with Decreasing Marginal Benefits of Information. Utrecht School of Economics, mimeo.
- Krackhardt, D. 1992. The Strength of Strong Ties: The Importance of Philos in Organizations. In: Nohria, N., and Eccles, R. (Eds.), *Networks and Organizations: Structure, Form, and Action*. Harvard Business School Press, pp. 216-239.
- Krebs, V., 2008. Consulting in Social Network Analysis [SNA] and Organizational Network Analysis [ONA]. Accessed 26 May, 2008. Available from <http://orgnet.com/consulting.html>.
- Leader Values - Getting Better Results Together. Accessed 26 May, 2008. Available from <http://www.leader-values.com>.
- Newman, M.E.J., Watts, D.J., and Strogatz, S.H., 2002. Random Graph Models of Social Networks. *Proceedings of the National Academy of Sciences USA* 99, 2566-2572.
- Padgett, J.F., and Ansell, C.K., 1993. Robust Action and the Rise of the Medici, 1400-1434. *American Journal of Sociology* 98(6), 1259-1319.
- Powell, W.W., Koput, K.W., and Smith-Doerr, L., 1996. Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in the Biotechnology. *Administrative Science Quarterly*, 41, 116-145.
- Raub, W., and Weesie, J., 1990. Reputation and Efficiency in Social Interaction: An Example of Network Effects. *American Journal of Sociology* 96, 626-654.
- Sabidussi, G., 1966. The Centrality Index of a Graph. *Psychometrika* 31(4), 581-603.
- Snijders, T.A.B., 1996. Stochastic Actor-Oriented Models for Network Change. *Journal of Mathematical Sociology* 21, 149-172.
- Snijders, T.A.B., 2001. The Statistical Evaluation of Social Network Dynamics. In: Sobel, M.E., Becker, M.P. (Eds.), *Sociological Methodology*. Basil Blackwell.
- Song S., Nerur S., and Teng J., 2007. An Exploratory Study on the Roles of Network Structure and Knowledge Processing Orientation in Work Unit Knowledge Management. *ACM SIGMIS Database*, 38, 8-26.
- Uzzi, B., 1996. The Sources and Consequences of Embeddedness for the Economic Performance of Organizations: The Network Effect. *American Sociological Review* 61(4), 674-698.
- Valente, T.W., and Foreman, R.K., 1998. Integration and Radiality: Measuring the Extent of an Individual's Connectedness and Reachability in a Network. *Social Networks*, 20, 89-105.
- Walker, G., Kogut, B., and Shan, W., 1997. Social Capital, Structural Holes and the Formation of an Industry Network. *Organization Science* 8(2), 109-125.
- Wasserman, S., Faust, K., 1994. *Social Network Analysis*. Cambridge University Press.
- Watts, D.J., and Strogatz, S.H., 1999. Collective Dynamics of Small World Networks. *Nature* 393, 440 - 442.

- Watts, D.J., 1999. Networks, Dynamics and the Small World Phenomenon. *American Journal of Sociology* 105(2), 493-527.
- Weidner, D., 2007. Special KM Topics – Social Network Analysis. Accessed 26 May, 2008. Available from <http://www.tmsasia.com>.
- Willer, R., 2007. The Role of Metanetworks in Network Evolution. *Journal of Mathematical Sociology* 31, 101-119.