Network Formation with Closeness Incentives

Berno Buechel

Abstract Closeness centrality is an index that has been widely used to assess the strength of an agent's position in a network of relationships. We study the formation of networks in a strategic setting, where every agent tries to optimize his closeness centrality. We investigate how the curvature of the benefit function (decreasing vs. increasing marginal returns) affects the set of stable networks and compare the results to the well-known connections model of Jackson and Wolinsky (JET 71, 1996). It turns out that our model can "replicate" the connections model in the sense that each result is translatable from one model into the other and the sets of stable networks coincide for certain specifications. We conclude that the two models incorporate the same kind of linking behavior and that grouping these "closeness-type" models means a first step in organizing network formation models by the type of incentives.

1 Introduction

Positions in social networks play a predominant role for economic outcomes. For example, consider a network of R&D collaborations in a technology-based industry. Companies which occupy a very "central" position are considered to better acquire and exploit knowledge that finally promotes their performance (e.g. [15]). In the field of social network analysis there is a long and rich history of studying benefits of network structures in various contexts. Beyond describing case studies, measures were developed that quantitatively assess the "merit" of certain network positions.[1]

B. Buechel
Institute of Mathematical Economics, Economic Behavior and Interaction Models
33501 Bielefeld, Germany
berno@wiwi.uni-bielefeld.de

[1] The underlying assumption of these approaches is that there are some structural features of networks (also called network statistics) that always have an impact on the opportunities or well-being of the agents, be it firms, people, or other units of decision.

A.K. Naimzada et al. (eds.), *Networks, Topology and Dynamics*, Lecture Notes in Economics and Mathematical Systems 613,

This paper considers the problem from a different perspective by asking how network structures can be affected by agents that strive for beneficial positions.[2] As the impetus of each individual's linking behavior we use one of the most customary centrality indices: closeness centrality [8]. Closeness captures the idea that it is beneficial for an agent to have short paths to many agents in the network. Applications range from performance of organizational units [17], over web-based communities, to status in school classes.[3]

We model a situation of two-sided link formation, based on the framework introduced by Jackson and Wolinsky [12]. An important example therein is the connections model they presented, which was intensively studied thereafter (see, e.g. [1]). The benefits of the connections model represent information transmission with some decay. As the decay is based on the length of communication paths, the benefits of the connections model are also considered as a "closeness-like" centrality index [3]. Since closeness and connections are clearly similar in spirit, the question arises, to what extent the results for the connections model persist when using closeness centrality instead.

Experience shows that in network formation models minor changes may have major effects. E.g. [6] and [11] both study network formation based on the same concept ("structural holes"), but use a different specification, which leads to quite different results. Based on these considerations this paper investigates robustness of the connections model: *How do the stable networks in the closeness model differ from those in the connections model and what is in general the significance of the curvature of the benefit function?* It turns out that the closeness model in its linear version "replicates" the results of the (symmetric) connections model, meaning that both payoff functions incorporate the same behavior.

In the next section we will introduce the model. Section 3 provides basic results. Section 4 compares the linear closeness model to the connections model.

2 Model

Before motivating and defining the closeness model, we introduce the necessary definitions.

2.1 Framework

Let $N = \{1, \ldots, n\}$ be a (finite, fixed) set of *agents/players*, with $n \geq 3$. A *network/graph* g is a set of unordered pairs, $\{i, j\}$ with $i, j \in N$. This set represents

[2] This idea can also be found in Rogers [16], who models the formation of weighted graphs using an index of social influence.

[3] Freeman [8] clarifies that closeness measures one dimension of centrality, while there are other dimensions, i.e. closeness does not sufficiently capture the intermediary role of some network positions.

who is linked to whom in a non-directed graph, i.e. $\{i, j\} = ij \in g$ means that player i and player j are linked under g. Let g^N be the set of all subsets of N of size two and G be the set of all possible graphs, $G = \{g : g \subseteq g^N\}$. A network can be seen as a (irreflexive, symmetric) binary relation on the player set. It can be represented by a matrix of zeros and ones called adjacency matrix.

Let $L_i(g)$ be the set of links that player i is involved in, that is $L_i(g) = \{ij \in g : j \in N\}$, and $l_i(g)$ its cardinality, called *degree*. An isolate is a player with degree zero and a pendant is a player with degree one (this structure is called a *loose end*).

A *circle* of size K is a sequence of K distinct players $(i_1, \ldots, i_K)$ such that $\{i_k, i_{k+1}\} \in g \ \forall k \in \{1, \ldots, K\}$, where $i_{K+1} := i_1$.

A *path* between two players i and j is a sequence of distinct players $(i_1, \ldots, i_K)$ such that $i_1 = i$, $i_K = j$, and $\{i_k, i_{k+1}\} \in g \ \forall k \in \{1, \ldots, K-1\}$. The (geodesic) *distance* between two players is the length of their shortest path(s), where the length is the number of links in the sequence. Formally, we can define the distance between two players ($d_{ij}(g)$) in a graph g by the corresponding adjacency matrix $A(g)$: $d_{ij}(g) := min\{k \in \mathbb{N} : A^k(g)_{ij} \geq 1; M\}$. If two players cannot reach one another (there does not exist a path connecting them), we define their distance as M, a number that is bigger than the feasible distances (see Sect. 2.2).

A graph is called *connected*, if there exists a path between any two players in the graph. A set of connected players is called *component*, if they cannot reach agents outside this set. A link is called *critical*, if its deletion increases the number of components in a graph. A graph is called *minimal*, if all links are critical. A *tree* is a connected network that is minimal.

Game Theoretic Framework

We base our model on a game-theoretic framework introduced by [14], [12] and [1]. Without defining the game explicitly, we take the "shortcut" of working with preferences and directly applying a stability concept.

For each player $i \in N$ a utility function represents his preferences over the set of possible graphs, $u_i : G \to \mathbb{R}$.

We work with the most basic equilibrium concept due to [12]: a network is considered as "stable" if no link will be added or cut (by two, respectively one player). Formally, a network g is pairwise stable (PS) or just *stable* if:

(1) $\forall ij \in g, \quad u_i(g) \geq u_i(g \backslash ij)$ and $u_j(g) \geq u_j(g \backslash ij)$
(2) $\forall ij \notin g, \quad u_i(g \cup ij) > u_i(g) \Rightarrow u_j(g \cup ij) < u_j(g)$

2.2 Closeness

The unique feature of the model presented here is the benefit function, which will be based on closeness centrality. Closeness incorporates the idea that an agent prefers networks, in which his average distance is short. Closeness is used in a wide variety of applications. We can identify the following basic arguments why closeness can be beneficial:

- *The higher your closeness, the smaller the distance to an arbitrary node in the network.* E.g. think of some researcher having a revolutionary idea. In the (co-author) network of this field of research, people with high closeness are likely to find out about the result much earlier.
- *The higher your closeness, the higher the expected spillovers from other agents.* As short geographic "distances" lead to external economies of scale, e.g. cars production in Detroit or Silicon Valley (electronics), it is plausible that short (network) distances have a similar effect, especially in times of well-developed information and communication technologies.
- *The higher your closeness, the higher your status.* In the friendship network of a school class one can assess the popularity of pupils by their closeness.
- *The higher your closeness, the better you can shape the community.* Networks with shorter paths facilitate quick diffusion of information/innovation. Agents with high closeness, therefore, can better spread their ideas.
- *The higher your closeness, the better you are informed.* The idea here is that accuracy of knowledge decreases with distances.

The listed arguments are not necessarily valid throughout all applications; the list only reflects how the importance of closeness can be justified. Our model is not dependent on the empirical validity of these motivations. The fact that some researchers or businessmen claim that closeness is desirable, provides enough justification to study network formation based on closeness incentives.

Definition of Closeness

We can generally define closeness such that benefits of an agent i gained by network g are decreasing with the (geodesic) distances of i to all other agents. To measure closeness there are some more details to look at.

To handle pairs that cannot reach one another, one can either restrict attention to the set of connected graphs, which would be a harsh assumption in a network formation game, or it must be defined what the distance of unconnected agents is. Here we define it as M. When i and j are connected, their distance is in $[1, n-1]$, hence let $M > n-1$.[4]

In the literature on centrality it is standard to normalize an index between 0 and 1. We follow this convention by defining *closeness* of an agent i as the following affine transformation of his average distance $\frac{\sum_{j \in N} d_{ij}(g)}{n-1}$:

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

[4] It is often convenient to define $M = n$. In this paper, however, we will keep it as a parameter.

There is another operationalization which is more prominent in the literature: the closeness definition according to Freeman [8]: $FrClose_i(g) := \frac{n-1}{\sum_{j \in N} d_{ij}(g)}$.[5] The author's trade-off here was: while Freeman's version (inverse distances) is much more customary, our closeness definition (reverse distances) more naturally separates the measurement of a structural feature of a network (network statistic) from its evaluation (by keeping the units, as argued in [18]). In the next subsection we will see that this choice does not restrict generality, e.g. if people strive just for Freeman-closeness, this is a special case of our model.

2.3 Model Specification

Our model is based on three major assumptions on individual behavior:

1. *The agents take linking decisions in respect to their degree and their closeness, where closeness is beneficial and links are costly.* To get a pure model we exclude all other aspects. One can think of any decision about adding or cutting links as a proposed exchange of average distance vs. degree: You can buy closeness by adding links; you can save costs by passing on closeness.
2. *The utility of a player is composed in an additive way by costs and benefits.* This assumption is not very restrictive as utility functions that are not additive separable may be transformed into this form. But it is a very convenient assumption: As the cross-derivatives are zero, the assumption uncouples the effects on utility coming from a change in closeness and a change in degree.
3. *The players are homogeneous in respect to preferences.* It is an interesting question to ask how networks evolve when players differ in their preferences (see, e.g. [9]). As applications of our model are very different in nature, however, we put emphasis on the different contexts that influence everybody's choice, not on the difference between agents (as also argued in [5]).

By introducing a (non-decreasing, twice differentiable) benefit function $b : [0,1] \to \mathbb{R}$ and a (non-decreasing, twice differentiable) cost function $c : [0, n-1] \to \mathbb{R}$, we can put all assumptions together to what we call the *closeness model*:

> All agents $i \in N$ decide about links according to preferences that can be represented by $u_i(g) = b(Close_i(g)) - c(l_i(g))$.

Although concave and convex cost functions are reasonable – concave costs represent the combination of fix costs and variable costs; convex costs represent the scarcity of resources (e.g. time) – we will restrict attention to linear cost functions

[5] In the original version Freeman closeness is only defined for connected graphs. The extension to all networks works with the definition of the distance of unconnected players (as M).

$c^{linear}(g) = \bar{c}l_i(g),\ \ where\ \ \bar{c} \in (0,\infty)$.[6] The justification is that a concave or convex cost function would induce similar behavior as the benefit function does when transformed with the inverse function. So these aspects are assumed to be absorbed by the benefit function.

For the benefit function we will distinguish three cases: concave shape, convex shape and linear shape.[7] The first one represents decreasing marginal returns. Formally, $\forall x, x', \Delta > 0$ a concave benefit function implies $b(x+\Delta) - b(x) \geq b(x'+\Delta) - b(x')$ whenever $x \leq x'$ (by the mean value theorem). Convexity implies increasing marginal returns: just let $x' \leq x$.

Remark 1. By taking the following convex benefit function $f(x) = [M - x(M-1)]^{-1}$, the benefits are equivalent to Freeman-closeness (with linear evaluation), because $f(Close_i(g)) = FrClose_i(g)$.

The *marginal costs* $\bar{c}$ are constant and serve as the parameter for our model. The marginal benefits depend on the network g and on the shape of the benefit function. Let $\beta_i^{ij}(g)$ denote the *marginal benefit* that link *ij* (either added or cut) means to player i in graph g. That is, $\beta_i^{ij}(g) := b(Close_i(g \cup ij)) - b(Close_i(g \setminus ij))$.

When players take linking decisions, they compare marginal costs and marginal benefits: in graph g player i is eager to form a link to j ($ij \notin g$) iff $\beta_i^{ij}(g) > \bar{c}$ and i wants to cut a link with k ($ik \in g$) iff $\beta_i^{ik}(g) < \bar{c}$.[8]

3 General Results

This section provides boundaries (thresholds of the parameter) for stable networks in the closeness model and addresses how they can be affected by the curvature of the benefit function.

3.1 Connectedness and Loose Ends

To have a shorter notation, we substitute two often needed units of closeness:

1. $T1 := \frac{1}{(n-1)(M-1)}$. This is the smallest possible change in closeness, as it corresponds to a shift in distance of 1. It occurs when two players, who were at distance two, form a link and only the distance between these two changes, e.g. because they are already directly linked to everybody else.

[6] In fact, this assumption restricts preferences to be quasi-linear in degree.

[7] In [10] the role of concave/convex benefits is nicely elaborated. [13] analyzes decreasing marginal returns in a similar model, but with one-sided link formation.

[8] When marginal benefits are equal to marginal costs, the player is indifferent. In this case he does not cut the link, respectively does not initiate the new link (but agrees when asked).

2. $T2 := \frac{1}{(n-1)}$. This is the change in closeness of a player that links with an isolate. As his distance shifts from M to 1, his closeness increases by $\frac{M-1}{(n-1)(M-1)} = T2$.

The following results provide two characteristics of all stable networks.

Proposition 1. *In a closeness model with linear costs and concave benefits the following holds:*

(1) If $\bar{c} < b(1) - b(1-T2)$*, all stable graphs are connected.*
(2) If $\bar{c} > b(T2) - b(0)$*, no stable graph exhibits loose ends.*

Proof. (1) Take any unconnected graph g. Take any player i and let $Close_i(g) =: x$. Linking with somebody of another component leads to a shift in closeness of at least $T2$. Because $x + T2 \leq 1$ and $b(\cdot)$ concave, it holds that $b(x+T2) - b(x) \geq b(1) - b(1-T2)$. By assumption the marginal costs are lower, such that i wants to form this link. As in any unconnected graph there exist two players who are not connected, they will alter the network structure, which makes g unstable.

(2) Take any network g with at least one pendant and let i be his (only) neighbour. Denote $Close_i(g) =: x$. Cutting the link to the pendant means a shift in closeness of $T2$. Because $x \geq T2$ and $b(\cdot)$ concave, it holds that $b(x) - b(x-T2) \geq b(T2) - b(0)$. By assumption the marginal costs are higher. Therefore, i will cut the link, which makes g unstable. $\square$

The intuition behind the result is that the thresholds of (1) and (2) are just the minimal and the maximal marginal benefit that a link to an isolated node can mean.[9]

If the benefit function is not concave but convex, these two thresholds just switch roles, as stated by the following proposition.

Proposition 2. *In a closeness model with linear costs and convex benefits the following holds:*

(1) If $\bar{c} < b(T2) - b(0)$*, all stable graphs are connected.*
(2) If $\bar{c} > b(1) - b(1-T2)$*, no stable graph exhibits loose ends.*

The proof is analogue to the proof of Proposition 1. Connectedness and non-existence of pendants heavily restrict the candidates for stable networks.

3.2 Existence

With the assumption of a convex benefit function, there is a very simple – admittedly not a very elegant – way of proving existence of stable graphs.

[9] I.e. the threshold in (2) is the marginal benefit of a new link in the empty graph $\beta_i^{ij}(g^{empty})$; and the threshold in (1) is the marginal benefit that cutting a link means to the center of a star $\beta_c^{ci}(g^{\star})$.

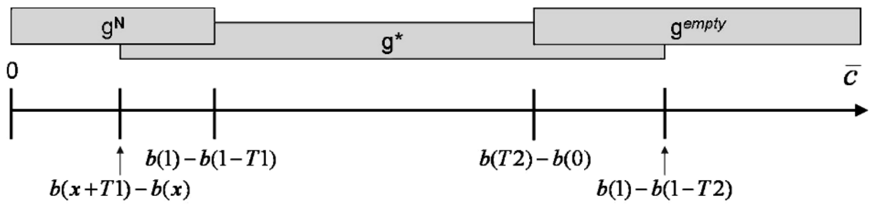

Fig. 1 Existence of stable networks for convex benefits

Proposition 3. *In a closeness model with linear costs and convex benefits the following holds: for any parameter value there exists at least one stable network.*

Proof. To show that for any marginal costs $\bar{c} \in (0,\infty)$ there exists a stable network, we take for low costs the complete graph, for high costs the empty network, and in the medium range the star. It is easy to verify that:

- The complete graph is stable if $\bar{c} \leq \beta_i^{ij}(g^N) = b(1) - b(1-T1)$. Remember that *T1* is the shift in closeness when distance increases by 1.
- The empty network is stable if $\bar{c} \geq \beta_i^{ij}(g^{empty}) = b(T2) - b(0)$.
- A star is stable if $b(x+T1) - b(x) \leq \bar{c} \leq \min\{b(1) - b(1-T2), b(x) - b(0)\}$, where $x := \frac{M}{M-1} - \frac{2n-3}{(M-1)(n-1)}$ is the closeness of a peripheral player (pendant). To verify the result, note that this condition precludes all possible deviations: (a) no peripheral players add a link $\bar{c} \geq b(x+T1) - b(x)$; and (b) the center does not cut a link $\bar{c} \leq b(1) - b(1-T2)$; and (c) no peripheral player cuts a link $\bar{c} \leq b(x) - b(0)$.

To prove existence for any marginal cost $\bar{c}$, it remains to show that (1) the lower bound of the star is below the upper bound of the complete network and (2) the upper bound of the star is above the lower bound of the empty network (see Fig. 1). 1. $b(x+T1) - b(x) \leq b(1) - b(1-T1)$ follows from $x + T1 \leq 1$ and convexity of $b(\cdot)$. And 2. $b(1) - b(1-T2) \geq b(T2) - b(0)$ follows from convexity of $b(\cdot)$; and $b(x) - b(0) \geq b(T2) - b(0)$ follows from $b(\cdot)$ increasing and $x \geq T2$. □

Figure 1 shows the idea of the proof: For any marginal cost, we can give a trivial example for a pairwise stable network.[10]

Remark 2. Figure 1 also contains the thresholds for Proposition 2 (on the right). In the case of concave benefits these two thresholds not only switch positions, but also switch their roles as stated in Proposition 1.

Besides these trivial examples (empty, complete, star) there are many more stable networks (which will be addressed in Sect. 4).

[10] For concave benefits the thresholds shift such that these trivial graphs do not span the whole parameter space. So in the case of concavity there are two "gaps" for which we could neither prove existence nor non-existence; for all other parameter values, existence is assured.

3.3 Pairwise Nash Stability

Besides pairwise stability there are other equilibrium concepts for network formation models, most of which are refinements of (PS). One of the most used stems from a non-cooperative framework and is called pairwise nash stability (PNS) (see, e.g. [2]). We can directly define it by just strengthening condition (1) of (PS): A network g is *pairwise nash stable* (PNS) if:

(1) $\forall i \in N, \forall l \subseteq L_i(g) \quad u_i(g) \geq u_i(g \backslash l)$.
(2) $\forall ij \notin g \quad u_i(g \cup ij) > u_i(g) \Rightarrow u_j(g \cup ij) < u_i(g)$.

In the closeness model, (PNS) is not always a proper refinement of (PS):

Proposition 4. *In a closeness model with linear costs and concave benefits the set of pairwise stable networks* $[PS]$ *and the set of pairwise nash stable networks* $[PNS]$ *coincide.*

One direction of the result follows directly from the definitions: $[PNS] \subseteq [PS]$. The other direction is more intriguing. Because of its length we omit the proof here and just present the main ideas[11]:

Calvó-Armengol and Ilkiliç [7] show that $[PNS]$ and $[PS]$ coincide, if the utility function $u(\cdot)$ satisfies a property called $\alpha - convexity$ in current links. Moreover, if costs and benefits are additively separable and marginal costs are constant, it is enough to show that the benefit function satisfies $\forall i \in N, \forall g \in G, \forall l \subseteq L_i(g)$,

$$\beta_i^l(g) \geq \sum_{ij \in l} \beta_i^{ij}(g), \tag{1}$$

where $\beta_i^l(g) := b_i(Close_i(g)) - b_i(Close_i(g \backslash l))$ denotes the marginal benefit that the deletion of the links (in l) means to some player i.

In essence, the condition says that the deletion of some of player i's links hurts him weakly more than the sequential deletion of these links, one at the time. For constant marginal costs it is intuitive that this is the condition requiring that deviations of cutting more than one link are only utility improving, if deviations of cutting just one link are, which is sufficient for $[PS] = [PNS]$.

To show that condition (1) holds in a closeness model with concave benefits, we need two steps: one step shows that the shift in closeness on the left-hand side of (1) cannot be smaller than the shift in closeness on the right-hand side. The other step exploits decreasing marginal returns (which guarantee, roughly, that multiple small reductions of closeness are not evaluated as severely as one big reduction).

The proof of Proposition 4 clarifies the role of the benefit function for the stability of networks: it is a genuine feature of the model that cutting one link at a time shifts closeness (weakly) less than cutting them at once. The concavity of the benefit function is just used to preserve this feature.

[11] The complete proof can be requested by the author.

This section showed how the curvature of the benefit function shifts thresholds for stable graphs. In the next section we study a special case, in which multiple thresholds coincide (to what we call "transition points").

4 The Linear Closeness Model

In the *linear closeness model*, we assume all players to have a linear cost function and a linear benefit function.[12] Without restriction of generality, we represent any player's preferences by $u_i^{linear}(g) = Close_i(g) - \bar{c}l_i(g)$. Note that by taking the id-function as benefit function, we mingle in this section what we distinguished before: the closeness of an agent and his benefit derived from closeness.

4.1 Transition Points

The first proposition is a corollary of Propositions 1 and 2, as the linear benefit function is a special case of both, concave and convex benefits functions.

Proposition 5. *Let again* $T2 := \frac{1}{(n-1)}$. *In the linear closeness model the following holds:*

(1) For $\bar{c} < T2$, *all stable graphs are connected.*
(2) For $\bar{c} > T2$, *no stable graph exhibits loose ends.*

Excluding pendants implies for the stable networks: (a) they cannot be minimal (i.e. a tree); (b) there exists at least one circle if the graph is non-empty; and (c) if the graph is connected, then it must contain at least n links. Observe that in this result two thresholds coincide: $b(1) - b(1 - T2) = b(T2) - b(0) = T2$. This is also true for the next transition point.

Proposition 6. *Let again* $T1 := \frac{1}{(n-1)(M-1)}$. *In the linear closeness model the following holds:*

(1) For $\bar{c} < T1$, *the unique stable network is the complete network.*
(2) $T1 \leq \bar{c} \leq T2$, *a star shaped graph is stable, but not necessarily unique.*

Proof. Remember that *T1* is the shift in closeness when distances shift by 1.

(1) The minimal increase in benefit that a new link can lead to for both its owners is $T1$; because a new link reduces at least the distance to the other player from 2 to 1. So, if costs are strictly lower than this, it follows immediately that nobody wants to cut a link in any graph (stability of complete graph) and any two players, who are not directly linked, will add a link (uniqueness).

(2) Shown in proof of Proposition 3 □

[12] As a consequence, the linear closeness model differs from Freeman-closeness (with linear evaluation), but it is equivalent to Freeman-Closeness with a certain concave benefit function.

Costs below *T1* are considered as very small; costs above *T2* are considered as very high. However, *T2* is not necessarily a threshold for uniqueness (of the empty graph): There is a third transition point (which can be bigger than *T2*).

Let *T3* be the maximal marginal benefit that a non-critical link can mean to both its owners. We claim that $T3 = \frac{n-1}{4(M-1)}$.[13] $T3$ occurs in the line graph, where the pendants form a link, respectively in a circle graph as the marginal benefit of cutting a link.

Proposition 7. *In the linear closeness model the following holds: (1) For* $\bar{c} > T3$ *every stable graph is minimal or empty. (2) If* $T3 \geq T2$,[14] *then for* $\bar{c} > T3$ *the unique stable graph is the empty network.*

Proof. (1) If a non-empty network is not minimal, then there must be at least one non-critical link. By the definition of *T3*, networks with such links cannot be stable in this cost range. (2) The empty graph is stable because $\bar{c} \geq T2$. For uniqueness note that any non-empty graph must contain either loose ends or circles. By Proposition 5 we can exclude all graphs with loose ends for $\bar{c} > T2$. By (1) we can exclude all graphs with circles for $\bar{c} > T3$. □

The transition points organize the equilibria in the parameter space. For very small costs and for very high costs, there are only trivial stable networks. In the medium cost range we can find a multitude of stable networks. Figure 2 shows one example of a stable network in the linear closeness model for *n=14*, *M=n* and medium costs.

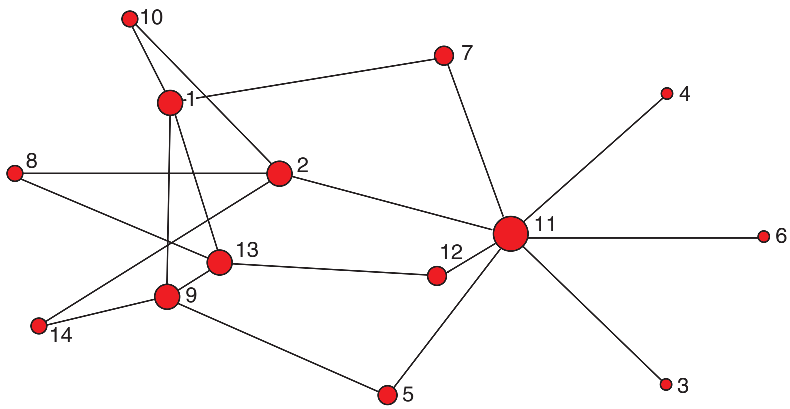

Fig. 2 Example of a stable network (the size and the position of a node indicates its closeness)

[13] The derivation of the value for $T3$ can be requested by the author.

[14] Mostly we will assume that M is such that $T3 \geq T2$ holds. We treat the exception of $T3 \leq T2$ in the next subsection as Proposition 8.

4.2 Comparison to Connections Model

In the famous example of the (symmetric) connections model, basically the following benefit is used: $Connections_i(g) = \sum_{j \in N \setminus i} \delta^{d_{ij}(g)}$, where $\delta \in (0,1)$.[15] So every reachable agent is of value, but this diminishes with distance. Like in the closeness model agents gain from short paths to other nodes. But there is also a difference: In the connections model agents benefit from having many nodes close to them; while in the closeness model agents benefit from having a small average distance.

While the motivation of the two models is similar, the results turn out to be almost identical.

Observe first that Propositions 5 and 6 correspond directly to the results of the connections model, where $T1 \hat{=} \delta - \delta^2$ and $T2 \hat{=} \delta$.

For n not too big, a computer can enumerate all networks and check for stability.[16] We did this for $n = 8$ with the connections model (taking $\delta = 0.5$ and $\delta = 0.8$), and for the closeness model once with the convex benefit function according to Freeman and once taking the linear closeness model (with $M = n$).

For $n = 8$ there are 12,346 different isomorphic graphs. In the linear closeness model only 45 of them are stable for some parameter range (greater than 0).[17] As depicted in Table 1, those 45 networks are not identical to the 63 stable networks with convex benefit function (Freeman), but overlap to some extent. The stable networks of the linear closeness model and the connections model overlap more heavily.

All of the above models are driven by similar linking behavior, which we call "closeness-type" incentives: there is high incentive to link to agents who are at high distance (or in another component) and there is low incentive to keep links that do not shorten some paths significantly. Interestingly, the differences within these models stem from specification details – be it increasing (instead of constant) marginal returns or level of decay – rather than from the choice of the model (connections

Table 1 Stable networks in the linear closeness model and related models for n=8

Number of stable networks (for some cost range)	Total	Also stable in linear closeness model
Freeman closeness	63	29
Connections $\delta = 0.5$	29	26
Connections $\delta = 0.8$	45	45

[15] By convention, here $M = \infty$ (see [12]).

[16] I thank Vincent Buskens for programming the routines to find all the stable networks for the various centrality measures.

[17] That is: we did not count the networks which are "stable" for only one point in the parameter space, e.g. the networks which are only stable if $\bar{c} = T1$.

vs. closeness). In this sense the connections model turns out to be "robust" and we conclude that the linear closeness model incorporates the same behavior as the connections model with some decay.

4.3 Trees

A very special case of the connections model occurs when the decay is very small or zero. Then distances become irrelevant and benefits only depend on the size of an agent's component. In this context the stable networks are trees, as non-critical links are worthless.

It turns out that the closeness model can also replicate this feature of the connections model by setting M sufficiently large[18]:

Proposition 8. *In the linear closeness model for marginal costs in the range* $T3 < \bar{c} < T2$ *the following holds:*

(1) All stable networks are trees.
(2) All trees are stable.

Proof. (1) Trees are characterized as minimal graphs that are connected. For $\bar{c} < T2$ all stable graphs are connected, as shown in Proposition 5(1). For $\bar{c} > T3$ all stable graphs are minimal as shown in Proposition 7(1).

(2) A graph is stable, if (a) nobody cuts a link and (b) no two players add a link. (a) As a tree is minimal, cutting a link leaves two components (unconnected groups of players). The more agents there are in the other component, the higher the loss of benefits. The highest incentive to cut is always given by the neighbor of a pendant; he loses closeness of $\frac{1}{n-1} = T2$. By assumption, marginal costs are lower than this (minimal marginal benefit), therefore no agent in a tree will cut a link. (b) Adding a link to a tree is an addition of a non-critical link (it is a property of trees to be maximally acyclic graphs). For $\bar{c} > T3$ this cannot be favorable for both (by the definition of *T3*). □

Remark 3. Note that many trees are also stable, when costs are below *T3* (see [4]).

This section showed that in the linear closeness model multiple thresholds (for stable networks) coincide. Given this feature, we can replicate the connections model with or without decay.

5 Concluding Remarks

We introduced a network formation model based on incentives to optimize closeness centrality. We analyzed how the set of stable networks depends on the curvature of the benefit function and compared the results to the connections model. It turns out

[18] Letting $M > \frac{1}{4}(n-1)^2 + 1$ assures that $T2 > T3$.

that the linear closeness model represents the same kind of behavior as the connections model, because each result is translatable from one model into the other and the sets of stable networks coincide for some parameter settings.

By grouping these "closeness-type" models, we have made a first step in organizing network formation models by the type of incentives. Accordingly, Buechel and Buskens [4] investigate the emerging networks in this setting and compare the results to incentives of a different type. Further research should clarify which structural patterns of networks emerge in which context and discuss the welfare implications.

The main limitations of our model – as in most models of strategic network formation – are the strong assumptions on behavior: Our agents are endowed with complete information and high rationality.

Acknowledgments I would like to thank Bastian Westbrock, Dinko Dimitrov, Andreas Brunhart, Tim Hellmann, Evan Shellshear and an anonymous referee for insightful comments. I gratefully acknowledge financial support from the international research training group "Economic Behavior and Interaction Models" funded by the German Science Foundation.

References

1. Bala V, Goyal S (2000) A noncooperative model of network formation. Econometrica 68:1181–1229
2. Bloch F, Jackson M (2006) Equilibrium definitions in network formation games. Int J Game Theory 34:305–318
3. Borgatti S, Everett M (2006) A graph-theoretic perspective on centrality. Soc Networks 28:466–484
4. Buechel B, Buskens V (2008) The dynamics of closeness and betweenness. Bielefeld University, Mimeo
5. Burger M, Buskens V (2006) Social context and network formation: experimental studies. ISCORE 234, Utrecht University
6. Buskens V, Van de Rijt A (2008) Dynamics of networks if everyone strives for structural holes. Am J Sociol (forthcoming)
7. Calvó-Armengol A, Ilkiliç R (2007) Pairwise-stability and nash equilibria in network formation games. Working paper (2007-01), Univ. Autnoma de Barcelona
8. Freeman L (1979) Centrality in social networks: conceptual clarification. Soc Networks 1(3):215–239
9. Galeotti A, Goyal S, Kamphorst J (2006) Network formation with heteregeneous players. Games Econ Behav 54:353–372
10. Goyal S, Joshi S (2006) Unequal connections. Int J Game Theory 34:319–349
11. Goyal S, Vega-Redondo F (2008) Structural holes in social networks. J Econ Theory (forthcoming)
12. Jackson M, Wolinsky A (1996) A strategic model of social and economic networks. J Econ Theory 71:44–74
13. Kamphorst J (2007) Network formation with decreasing marginal benefits of information. Utrecht School of Economics, Mimeo
14. Myerson RB (1991) Game theory: analysis of conflict. Cambridge University Press, Cambridge

15. Powell WW, Koput KW, Smith-Doerr L (1996) Interorganizational collaboration and the locus of innovation: networks of learning in the biotechnology. Adm Sci Q 41:116–145
16. Rogers B (2006) A strategic model of network status. Working paper (2006-01), Caltech, Pasadena
17. Song S, Nerur S, 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
18. Valente TW, Foreman RK (1998) Integration and radiality: measuring the extent of an individual's connectedness and reachability in a network. Soc Networks 20:89–105