UNIVERSITY OF COPENHAGEN

Assortative matching with network spillovers

Bjerre-nielsen, Andreas

Published in:
Journal of Economic Theory

DOI:
10.1016/j.jet.2020.105028

Publication date:
2020

Document version
Peer reviewed version

Document license:
Unspecified

Citation for published version (APA):
Bjerre-nielsen, A. (2020). Assortative matching with network spillovers. *Journal of Economic Theory*, *187*, [105028]. https://doi.org/10.1016/j.jet.2020.105028

Assortative matching with network spillovers.[1]

Andreas Bjerre-Nielsen[2]

University of Copenhagen, Øster Farimagsgade 5, 1353 Copenhagen K., Denmark

Abstract

This paper investigates endogenous network formation by heterogeneous agents. The agents' types determine the value of linking and we incorporate spillovers as utility from indirect connections. We provide sufficient conditions for a class of networks with sorting to be stable for low to moderate spillovers; with only two types these networks are the unique pairwise stable ones. We also show that this sorting is suboptimal for moderate to high spillovers despite otherwise obeying the conditions for sorting in Becker (1973). This shows that in our sorted networks a tension between stability and efficiency is present. We analyze a policy tool to mitigate suboptimal sorting.

Keywords: network formation, under-connectivity, assortative matching, network externalities, one-sided matching.

JEL classification: C71; C78; D61; D62; D85.

1. Introduction

Social relations and their network structures are fundamental in almost all aspects of our lives: which jobs we get, how we perceive the world, the decisions we make, etc. (Jackson, 2019). A ubiquitous finding in studies of social relations is the tendency to form more ties with people similar to one-self, i.e. the pattern known

[1]This paper has grown out of Bjerre-Nielsen (2015). The author is grateful to Peter Norman Sørensen who supervised his PhD. A big thanks to Jan Eeckhout, Matthew Jackson, John Kennes, Jesper Rüdiger, Hans Keiding, Bartosz Redlicki, Thomas Jensen, David Dreyer Lassen and seminar participants at CTN 2018, CoED 2015, CoopMAS 2015, EDGE 2015, University of Copenhagen, Stanford University and Universitat Pompeu Fabra. This paper is a revised version of a chapter in the author's PhD dissertation at University of Copenhagen from 2016. The work was funded by University of Copenhagen.

[2]University of Copenhagen, email: andreas.bjerre-nielsen@econ.ku.dk.

as *sorting* or homophily, see the meta-study McPherson et al. (2001). Pioneered by Becker (1973), economic research has contributed to the understanding of sorting by providing mathematically sufficient conditions for sorting in marriage- and labor markets to be stable and optimal. The essential condition for sorting is *supermodularity*. This condition entails complementarity in type is such that similar types gain higher value from linking than dissimilar ones. Previous research has not analyzed what implications supermodularity has for assortative matching in the context of networks, either with or without utility from indirect connections.

We extend assortative matching to the context of networks. We demonstrate fundamental properties for a class of networks which we label as having *sorted connectivity*. We require these networks to satisfy the following conditions: *perfect sorting* such that all agents link only with agents of the same type; *type-connectivity* whereby any two agents of the same type are connected; and, *no-link surplus* whereby all available links are used. Intuitively, these networks maximize utility both at the individual and aggregate levels, conditional on perfect sorting. We show that networks with sorted connectivity exist when the number of agents for each type exceeds the number of links allowed per agent and a regularity condition holds, see Proposition 3.

We provide a novel, parsimonious framework which unites the frameworks of complementarity between heterogeneous agents (Becker, 1973) and utility from indirect connections to friends of friends (Jackson and Wolinsky, 1996). In our setup every agent has, at most, a fixed number of links, which reflects limited amounts of time and effort.[3] We investigate stability of networks in the following sense: no two agents can form and/or delete links in the network to improve their joint payoff (we allow for transfers of utility). Our central result is that if Becker's condition for supermodularity holds and spillovers (i.e magnitude of utility from indirect connections) are low to moderate, then networks with sorted connectivity are stable; with only two types of agents it further holds that every stable network has sorted connectivity. See details in Theorem 1.

We also investigate the efficiency for networks with sorted connectivity. For moderate to high spillovers and supermodularity we show that networks with sorted connectivity are inefficient. Therefore, we may refer to inefficient sorted networks as 'under-connected' as they have too little connectivity across types relative to the efficient networks. In the situation where there are only two types of agents, and

[3]A limited number of partners is consistent with empirical research: Ugander et al. (2011) show that the average number of social ties for the entire Facebook network is a few hundred, and likewise Miritello et al. (2013) show the number of phone calls for millions of people is also limited.

supermodularity holds, we strengthen the results: for low spillovers, networks with sorted connectivity are efficient – otherwise, networks with connectivity across types are efficient. Note that among these results, the latter is about global efficiency whereas the former is about comparative (relative) efficiency, see Theorem 2 for details.

The complementarity captured by supermodularity resembles situations where more similarity in matches increases the joint utility. In these situations heterogeneity in type may refer to productive and non-productive capabilities as well as other characteristics with synergy between types. Natural examples include geography or language, as suggested by Church and King (1993). Matching contrary to these characteristics may lead to increased transaction costs or miscommunication and, therefore, to lower aggregate productivity. Another example is combined effects in skills – it could be that matching workers or students with similar skills results in higher joint utility, e.g. as in the classic O-ring model of Kremer (1993). A property of this complementarity is exogeneity; synergy depends only on agents' pre-defined types, not other parts of the matching/network. We note there are other kinds of homophily/sorting on exogenous characteristics, e.g. eye-color, height. However, these kinds of homophily are often based directly on mutual preference for similarity and do not require conditions on the joint utility. This difference often discussed as transferable versus non-transferable utility, where this paper uses the former framework.

Our results show that Becker's supermodularity condition is no longer sufficient for sorting to be either stable or efficient. This comes from the fact that agents face a trade-off between two sources of utility: on the one hand, complementarity implies that increased sorting leads to higher direct utility as links between similar agents are more valuable; on the other hand, positive spillovers entail that more sorting can lead to a loss of utility from indirect connections. Therefore, instability and inefficiency of sorting stems from the utility of indirect connections dominating complementarity, which imply that the stable networks are under-connected. However, when combined, our results show that if there are moderate spillovers and supermodularity, then networks with sorted connectivity are stable but inefficient. The reason is that the underlying thresholds governing the two properties are not identical. The intuition behind this incompatibility of stability and efficiency is that two agents forming and removing links will internalize the direct utility from sorting, but they do not internalize the utility from indirect connections for third party agents. We show by visual inspection that the scope for incompatibility (i.e. the region of 'moderate' spillovers) widens as the number of agents grows and strength of complementarity increases. We demonstrate in Proposition 4 how to enact policies that curb excessive sorting by leveraging contracts that make payoffs conditional on links. These new insights can help policymakers within organizations to design better internal networks by overcoming under-connectivity, e.g. between individuals within schools, corporations or organizations.

The policy implications are easy to see in a stylized example. Suppose there are two islands with costly transportation from one to the other. If spillovers are moderate in the sense of our model, then the network is under-connected – no inhabitants of either island want to establish connections with inhabitants of the other island as their individual payoff is too small. Nevertheless, by paying agents to connect, everyone could be better off. That is, policies that foster connectivity across the islands can increase the efficiency of the underlying network. With more islands that are all disconnected, the problem can compound and thus the scope for policies increases. Although unrealistic, the results should translate to situations with high levels of homophily combined with strong complementarity and/or many agents. For instance, new empirical work has already shown that a few connections between otherwise connected sub-communities in online social media can foster diffusion of information that otherwise would be unlikely (Park et al., 2018). Our model also helps us to understand the potential consequences of forming organizations that consist of disjoint parts, e.g. school classes, company divisions. If the organizational parts are not encouraged to interact across affiliated parts, it may lead to no interaction (which is suboptimal), as the disjoint structure provides an implicit complementarity among members from the same part of the organization.

This paper also makes a number of additional contributions towards the understanding of assortative matching. We establish that supermodularity is sufficient for stable networks to contain a general pattern of sorting by type without utility from indirect connections, see Proposition 1. We also show that if the agent population is very large, then sorting is the unique strongly stable outcome (i.e. the core), when Becker's complementarity condition holds and spillovers are not too high, see Proposition 5.[4]

Literature. In what follows we review relevant literature and discuss our results in the context of the most related work. A seminal mathematical work on sorting and segregation is Schelling (1969, 1971). Although related, the modeling differences between Schelling's spatial model and networks (or matchings) are stark; networks are more flexible and allow for connections between any individuals and utility from indirect connections. Two-sided matching captures agents from two distinct sides

[4]The level of spillovers satisfies asymptotic independence in social connections.

who match, e.g. in labor and dating markets. The inaugural study on sorting is Becker (1973), which has since been extended to a search setting with match frictions, see Shimer and Smith (2000) and a recent review in Chade et al. (2017). Most of the earlier work on sorting investigates two-sided matching but we use a one-sided matching model. The research on one-sided assortative matching has been limited to formation of partnerships and clubs, which correspond respectively to one-to-one and many-to-many matchings (Farrell and Scotchmer, 1988; Kremer, 1993; Durlauf and Seshadri, 2003; Legros and Newman, 2002; Pycia, 2012; Baccara and Yariv, 2013; Xing, 2016).[5] All the research on one-sided assortative matching finds conditions for sorting which correspond to type complementarity in Becker (1973). Yet, none of the above papers allows for general linking beyond partnerships and clubs, or considers network spillovers. We relax both of these assumptions. We have carefully chosen our framework to use the fundamental concepts from the earlier literature, i.e. supermodularity, a finite capacity for forming links, and pairwise link formation. Our main contribution to the literature on assortative matching is to show which extra conditions in combination with supermodularity lead to stability and efficiency in networks with utility from indirect connections. Another key contribution is that we show potential incompatibility between stability and efficiency and that policies that create incentives to link can fix the issue. We also extend the framework of assortative matching without externalities to a one-sided setting with many partners but without restrictions on link structure (i.e. limited to clubs): Proposition 1 establishes that sorting is stable using a novel measure of sorting which is tractable in equilibrium.

The only paper that investigates sorting in networks with externalities is de Martí and Zenou (2017); they also model type complementarity and positive spillovers.[6] Their results show the existence of sorted networks that are stable yet inefficient due to the lack of linking across types.[7] Although this model is similar to ours, there are crucial differences that motivate our analysis. The essential difference is that we use exogenous complementarity which is independent of the network, while de Martí and Zenou (2017) use endogenous complementarity. de Martí and Zenou (2017) specifically assume that complementarity is strongest when the level of sorting is high, while there is no complementarity when there is negative sorting (i.e.

[5]Buchanan (1965) defines clubs as groups where one's utility depends on all other members. This means that clubs are networks with the implicit assumption that any agents of the same group are all linked. In the networks literature, such groups are known as cliques.

[6]Note that this paper was developed independently and without awareness of de Martí and Zenou (2017).

[7]See Propositions 1.ii, 4.iii in de Martí and Zenou (2017) for results on stable sorting; their Proposition 6 contains results on inefficiency.

tendency to link with dissimilar types). This different assumption makes our results considerably stronger than de Martí and Zenou (2017) when there are two types: we establish that sorting constitutes the *unique* set of stable networks; we show there exists a *globally* efficient network that is sorted but has connectivity between groups and we demonstrate that this network is *implementable* through a simple policy, see Proposition 4. In addition, our results also apply more generally as they are neither limited to only five agents of each type, nor to sub-structures of within-type networks being either stars or cliques, nor to only two types.

The essential difference between how we model complementarity here and how de Martí and Zenou (2017) model it also implies that we interpret our results differently. First, our results are relevant in cases when we do *not* expect endogenous complementarity (see example above). Second, the modeling of complementarity also entails that the source of inefficiency is different. In our setup suboptimality stems from misaligned incentives which entail there is a general incompatibility of efficiency and stability of as in Jackson and Wolinsky (1996). Thus, as sorting is the unique stable outcome, no one wants to volunteer to build the bridge between communities which increases overall welfare unless there is an outside readjustment of the incentives, e.g. by policy as we explore or by allowing different contracts (Bloch and Jackson, 2007). On the contrary, in de Martí and Zenou (2017) the network with two connected communities, is both efficient and stable. Thus, suboptimality is not due to incompatibility of efficiency and stability, but rather that agents ended up in one of the pairwise stable networks characterized by low welfare.[8]

There is a vast literature on optimal networks under externalities. The field has a long tradition and begins with the general formulation under the quadratic assignment problem (Koopmans and Beckmann, 1957). The field of matching and networks under externalities was revolutionized by Jackson and Wolinsky (1996) who demonstrate that there is an incompatibility in networks between stability and efficiency; Bloch and Jackson (2007) extend these results to show that the tension is preserved when allowing for more coordination and more flexible transfers between agents. Although Jackson and Wolinsky (1996) as well as Bloch and Jackson (2007) show that the incompatibility between stability and efficiency holds generally, they provide very little in terms of what structure inefficient networks can have. Our

[8]The results in de Martí and Zenou (2017) do not rule out that there can exist stable networks (e.g. some amount of connectivity between groups) that are more efficient than a sorted network. This follows as de Martí and Zenou (2017) have multiplicity in equilibria and they only establish relative inefficiency between two networks (complete network and perfectly sorted network of cliques).

contribution is to provide explicit structure to the incompatibility in the context of sorting. This extension may seem small but it has important implications - sorting is a fundamental pattern in empirical networks (McPherson et al., 2001) and therefore we show that the incompatibility may hold widely.

The most relevant research on exogenous complementarity in networks is Johnson and Gilles (2000); Jackson and Rogers (2005); Galeotti et al. (2006); the first assumes agents all have a unique type with linking costs proportional to their distance, while the two latter use an islands type of model (where agents have same type). Johnson and Gilles (2000) shows existence of a pairwise stable equilibrium with local connectivity between adjacent types, possibly with local cliques where all agents within a given range are connected. Jackson and Rogers (2005) shows that clustering and short paths are robust features among both pairwise stable networks and efficient networks with full linkage among same type. Galeotti et al. (2006) investigate minimally connected networks in a setup with one-sided link formation.[9]

There are other strands of literature on homophily in network formation e.g. Currarini et al. (2009, 2010) and Bramoullé et al. (2012). Their approach, however, is different: we use a one period model with strategic link formation, while they rely on matching sequences that are dynamic and stochastic. Currarini et al. (2009, 2010) investigate how differences in community sizes play a role in explaining empirical phenomena, including homophily. Bramoullé et al. (2012) investigate the conditions for long run integration of a network. Other literature has investigated the role of homophily in a model combining referral networks and a labor market (Montgomery, 1991; Galenianos, 2018).

Paper organization. The paper proceeds as follows: Section 2 introduces the model; Section 3 investigates sorting under no externalities; Section 4 analyzes the setting with externalities, focusing on sorting and its potential suboptimality, and; Section 5 concludes with a discussion of assumptions. All proofs are found in Appendix Appendix A.

2. Model

Let $N = \{1, .., n\}$ constitute a set of agents. Each agent $i \in N$ is endowed with a fixed measure of *type*, $x_i \in X$, where $X \subset \mathbb{R}$ is the set of (realized) types for agents in N. Let $\bar{x} = \max X$ and $\underline{x} = \min X$. Define the vector of types $\mathcal{X} = (x_1, x_2, ..., x_n)$.

[9] Note that one-sided link formation is based on the setup of Bala and Goyal (2000) which only requires the weaker equilibrium concept, Nash stability, as links do not need mutual acceptance.

Let the agents' type be sorted in descending order according to their label such that $x_l \geq x_{l+1}$ for $l = 1, .., n-1$.

Linking and networks. Two agents $i, j \in N$ may *link* if they both accept it. Any of the two agents who link may break the link without mutual consent. A link between i and j is denoted $ij \in \mu$, where the set μ consists of links and is called a *network*. The set of all networks is denoted $M = \{\mu | \, \mu \subseteq \mu^c\}$, where μ^c is the *complete* network in which all agents are linked.

A *coalition* of agents, t, is a subset of agents (i.e. $t \subseteq N$) such that $t \in T$, where T is the power set of N excluding the empty set. For a given group, t, define $\mathcal{X}(t)$ as the vector of types in descending order over each of the agents in t. A *coalitional move* is a set of actions implemented by a coalition that moves the network from one state to another. A move from μ to $\tilde{\mu}$ is *feasible* for coalition t if added links, $\tilde{\mu} \backslash \mu$, are only formed between members of coalition t and deleted links, $\mu \backslash \tilde{\mu}$, only contain members of coalition t.

Network measures. The *neighborhood*, ν, is the set of agents who an agent links to: $\nu_i(\mu) = \{j \in N : \, ij \in \mu\}$. The number of neighbors is called *degree* and denoted $k_i(\cdot)$ for i. A *path* is a subset of links $\{i_1 i_2, i_2 i_3, ..., i_{l-1} i_l\} \subseteq \mu$ where no agent is reached more than once; the *length* of a path is the number of links in its set. The *distance* between two agents, i, j, in a network is the length of the shortest path between them - this is denoted $p_{ij} : M \to \mathbb{N}_0$; when no path exists then the distance is infinite.

Utility. The utility accruing to agent i is denoted u_i. An agent's utility equals benefits less costs, expressed mathematically as $u_i = b_i - c_i$. The aggregate utility is denoted $U(\cdot)$. We model costs of linking indirectly through an opportunity cost of linking. We do this through a *(degree) quota* on links, κ, which is the maximum number of links for any agent, i.e. for $i \in N$, it holds $k_i(\cdot) \leq \kappa$. We say there is *no linking surplus* when all agents have a degree equal to the degree quota, i.e. $\forall i \in N : k_i = \kappa$. The benefit to agent i is a weighted sum consisting of two elements; network and individual value:

$$b_i(\mu) = \Sigma_{j \neq i} w_{ij}(\mu) \cdot z_{ij}. \tag{1}$$

The network factor $w_{ij}(\mu)$ is a function of network distance. The individual link value is z_{ij}, which measures the personal value to i of linking to j - the value is a function of the two partners' types $z_{ij} = z(x_i, x_j)$. The function z is assumed twice differentiable as well as taking positive and bounded values.[10] Let the *total link value* be defined as the value of linking for the pair, i.e. $Z_{ij} = z_{ij} + z_{ji}$.

In order to derive results, a restriction of payoffs is necessary. The essential feature of the total link value for sorting is complementarity in type:[11]

Definition 1. *The link value has **supermodularity** if* $\frac{\partial^2}{\partial x \partial y} Z(x, y) > 0$. *This entails:*

$$Z(x, \tilde{x}) + Z(y, \tilde{y}) > Z(x, y) + Z(\tilde{x}, \tilde{y}), \qquad x > \tilde{y},\, \tilde{x} > y. \tag{2}$$

The network components are further restricted in the analysis under externalities in Section 4.

The game framework. This paper explores a static setting of one period. Agents' information about the payoffs of other agents is complete. Together the players, action, utility and information constitute a game.

We assume that any pair of agents can transfer 'utility' between them. Let a *net-transfer* from agent j to agent i be denoted as $\tau_{ij} \in \mathbb{R}$ such that $\tau_{ij} = -\tau_{ji}$, which implies non-wastefulness of utility. The matrix of net transfers is denoted τ. Transfers can be exchanged by any pair of agents. We specifically assume that for any pair of agents there is mutual dependence between transfers and their link, if they have one. This entails that a transfer cannot be changed unless both agents agree, otherwise the non-consenting agent can break the link. Conversely, if the link is broken without mutual consent, then the transfers are set to zero. Although this seems similar to Bloch and Jackson (2007), the conditionality here is only between agents who are linked.

Stability. We define network stability using coalition moves. A coalition t is *blocking* a network μ with net-transfers τ if there is a feasible coalition move from network μ to network $\tilde{\mu}$ with $\tilde{\tau}$ where all members in t have a higher net-payoff after the move.

We employ two concepts of stability. The first is *strong stability*: this is satisfied for a network if there exist transfers such that no coalition (of any size) may have a feasible move that is profitable for all its members. The second concept, *pairwise (Nash) stability*,[12] is similar but has weaker requirements: it holds when there exist transfers where it holds that no coalitions of at most two agents may block. A further discussion of the stability concepts is found in Section 5.

[10] The upper bound rules out an infinite number of links in equilibrium.

[11] Complementarity between type corresponds to cheaper links between same/similar types used in the models of Johnson and Gilles (2000); Jackson and Rogers (2005); Galeotti et al. (2006).

[12] This is also known as bilateral stability, cf. Goyal and Vega-Redondo (2007).

Our pairwise definition of stability is stricter than that of Jackson and Wolinsky (1996). However, the stricter requirement enables substitution of links (simultaneous deletion and formation), which is a necessary requirement for establishing results in the matching literature.

A noteworthy feature is that strong stability implies pairwise stability; thus any condition valid for all pairwise stable networks also applies to any strongly stable network. In addition, without utility from indirect connections (i.e. no spillovers), every pairwise stable network is also strongly stable, see Lemma 1. Note also that any strongly stable network requires efficiency (coalition of all agents can implement any network). We use the efficiency property of strongly stable networks to derive the structure of these networks in Proposition 5.

3. Analysis: no spillovers

This brief section analyzes the setting without utility from indirect connections. We begin with defining our measure of sorting. The concept of sorting that we employ is a generalization of the sorting when there is a single partner, such as in Becker's marriage market. The shape of sorting is such that a high type agent has partners which weakly dominate in type when compared partner-by-partner with the partners of a lower type agent. Note the comparison is done over the sorted set of partners type $\mathcal{X}$. The sorting pattern is mathematically defined as:

Definition 2. ***Sorting in type*** *holds in μ if for all pairs i, j such that $x_i > x_j$ it holds that:*

$$\mathcal{X}(\nu_i(\mu)/\{j\})_l \geq \mathcal{X}(\nu_j(\mu)/\{i\})_{l+l^*}, \qquad \forall l \in \{1, .., k^*\},$$

where $k^* = \min(k_i(\mu), k_j(\mu))$ and $l^* = \max(k_j(\mu) - k_i(\mu), 0)$.

Our first result is that sorting in type emerges under the same conditions as in Becker (1973) when network externalities are absent:

Proposition 1. *If there is supermodularity and no externalities, then for any pairwise stable network there is sorting in type.*

The proof of this proposition follows by establishing that a pairwise stable network must be strongly stable without externalities; then we use that strongly stable networks are efficient and show that sorting in type must hold under efficiency.[13]

[13] Note that the current proof relies on comparing partner order of l for i with order $l + l^*$ for j in Definition 2. We conjecture that this can be relaxed to comparison partner of order l of i with order $l + l^*$ of j.

4. Analysis: positive spillovers

We proceed to a more general context where indirect connections matter for utility. Whenever we allow for externalities we restrict our attention to two forms of linking utility.

$$w_{ij}(\mu) = \begin{cases} \delta^{p_{ij}(\mu)-1}, & \text{constant decay,} \\ \mathbf{1}_{=1}(p_{ij}(\mu)) + \delta \cdot \mathbf{1}_{\in[2,\infty)}(p_{ij}(\mu)), & \text{hyperbolic decay,} \end{cases} \tag{3}$$

where $\mathbf{1}_{\in(1,\infty)}(l)$ is the Dirac measure/indicator function of whether $1 < l < \infty$.

The first and more general setting is where utility from connections decays over increasing distance at a constant exponential rate. This corresponds to benefits from linking in the 'connections-model' from Jackson and Wolinsky (1996). The other case is when externalities from indirect connections are discounted equally at any distance if there is a connection, i.e. a finite path length. This second case is referred to as hyperbolic decay and entails that there is no decay beyond that from distance one (linked) to distance two.

The introduction of externalities to our framework implies that the pairwise utility no longer depends only on the total link benefits. As a consequence, sorting is not guaranteed to be either stable or efficient. The intuition for this is straight forward: externalities entail that the total welfare from sorting is internalized for the pair, while the total welfare for indirect connections are not internalized. We see this by inspecting the utility functions. Suppose that $\hat{g}$ is a pairwise deviation such that agents i, j form a link. Then the pairwise total net utility from deviation can be expressed as follows under externalities:

$$u_i(\hat{g}) + u_j(\hat{g}) - u_i(g) - u_j(g) = Z_{ij} + \sum_{k\in\{i,j\},\, l\notin N\setminus\{i,j\}} z_{kl} \cdot \left(\delta^{p_{kl}(\hat{g})} - \delta^{p_{kl}(g)}\right) \tag{4}$$

From the analysis in the previous section we found that, in the absence of externalities, sorting prevails. In the above equation this incentive to sort is captured by the component Z_{ij}. Therefore, we see that the total benefits from sorting are preserved for the pair.

The total benefits to all agents that accrue from agents i, j forming a link are:

$$\begin{aligned} U(\hat{\mu}) - U(\mu) &= u_i(\hat{\mu}) + u_j(\hat{\mu}) - u_i(\mu) - u_j(\mu) + \sum_{l\notin N\setminus\{i,j\}} [u_l(\hat{\mu}) - u_l(\mu)] \\ &= u_i(\hat{\mu}) + u_j(\hat{\mu}) - u_i(\mu) - u_j(\mu) + \sum_{l\notin N\setminus\{i,j\},\, l'\in N,\, l'\neq l} Z_{ll'} \cdot (w_{ll'}(\hat{\mu}) - w_{ll'}(\mu)) \end{aligned} \tag{5}$$

Inspection of Equation 5 informs us that the pairwise utility of linking does not capture the aggregate gains from linking. Moreover, we see that the gains not captured correspond to the indirect benefits that others receive from the deviation. This implies that there is a disparity between the pairwise incentives and total welfare: the pairwise incentives capture the full benefits of sorting but not the full gains from lower distances between agents.

4.1. Finite population

We begin with the situation where there are a finite number of agents. Before starting the analysis of networks under externalities we define some useful concepts. Naturally, we call networks perfectly sorted when the agent of each type only link among themselves. A further important distinction is whether the subnetworks for each type are connected among agents of the same type. Such connectivity is sufficient for agents to reap all the gains of utility from indirect connections when sorting occurs if there is hyperbolic decay in spillovers. Finally, we want to ensure that there is no surplus of links as this would imply wastefulness, which is not in the interest of agents as they always benefit from linking. Combining these concepts we can introduce our main concept, sorted connectivity of networks:

Definition 3. *A network has **perfect sorting** if every linked pair of agents have the same type.*

Definition 4. *A network is **type connected** if every two agents of the same type is connected.*

Definition 5. *An agent i has **link-surplus** in a given network if i's number of links is lower than the degree quota.*

Definition 6. *A network is **sort-connected** if the network (i) is perfectly sorted, (ii) is type connected and (iii) no agents have a link surplus in the network.*

We now turn to type self-sufficiency, which requires that there is potential for each type to perfectly sort and have no link surplus. This concept is important and plays a critical role for the existence of sort-connected networks.

Definition 7. *A type of agents $x \in X$ is **self-sufficient** if $n_x > \kappa$.*

We briefly investigate the situation when type self-sufficiency does not hold, i.e. $n_x \leq \kappa$ for one or more types $x \in X$. This is seen from the following statement:

Proposition 2. *Suppose there is supermodularity and there is not type self-sufficiency:*

(i) if $n \leq \kappa + 1$ then the complete network is the only pairwise stable network and the only efficient network;

(ii) if $n > \kappa + 1$ and there are two types where $n_{\bar{x}} = n_{\underline{x}}$, then every network where every agent has $n_{\bar{x}} - 1$ same-type links, and $\kappa - n_{\bar{x}} + 1$ cross-type links is stable and efficient.

We move on to examining sorted connectivity in networks. We note that the remainder of this subsection is restricted in two ways. First, by confining our analysis to the setting where there is self-sufficiency for each type. Second, we exclusively focus on the case of hyperbolic decay as it provides for more intuitive and more immediate results without restrictions on the network. As noted earlier, a more general exposition is found in Supplementary Appendix Appendix B.

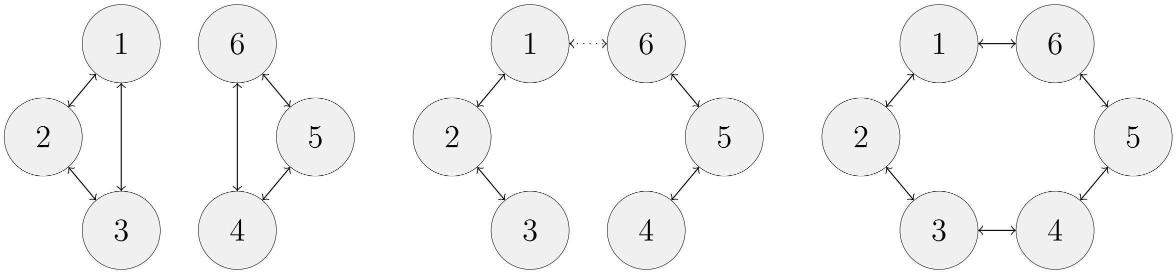
The aim is to show that sort-connected networks are stable when the strength of utility from indirect connections is low to moderate. Moreover, when the strength of utility from indirect connections is moderate to high, the networks are suboptimal despite fulfilling Becker's complementarity condition, i.e. supermodularity. We will see that the suboptimality arises because the network is under-connected relative to the efficient network.

We begin our analysis by presenting an illustration of the situation. We want to show that for strength of utility from indirect connections below the threshold, δ^{stab}, any network with sorted connectivity is also pairwise stable. Moreover, we will show that there exist networks with higher aggregate utility when utility exceeds another threshold, δ^{opt}. In order to motivate and capture the intuition we provide simplified results in Example 1. The example is based on hyperbolic decay of network externalities, however, it can be easily adapted to constant decay.[14] The example is graphically represented in Figure 1.

Example 1. *There are six agents; three of high type (1,2,3) and three of low type (4,5,6). Moreover, there is supermodularity, degree quota of two (κ=2) and hyperbolic decay. Define two networks: a network with sorted connectivity, $\mu = \{12, 13, 23, 45, 46, 56\}$, see Figure 1.A; a sort-connected network with bridges (see Definition 8), which we denote as $\tilde{\mu} = \{12, 23, 34, 45, 56, 61\}$, see Figure 1.C. We show in this example that, for a range of decay-factors, μ is pairwise stable, yet suboptimal. In this setup there is a unique move which is both feasible and payoff relevant.[15] This move consists in*

[14] In an old paper version, Bjerre-Nielsen (2015), we compute the example under constant decay of spillovers.

[15] Under pairwise stability at most one link can be formed in a single move. As the value of every



(A): Segregated network, μ. Pairwise stable if $\delta \leq \delta^{stab}(\hat{Z})$.

(B): Pairwise move by agent 1, 6 who form a link together and remove links 13, 46.

(C): Connected network, $\tilde{\mu}$. Efficient if $\delta \geq \delta^{opt}(\hat{Z})$.

Figure 1: *Sorted network is stable but inefficient.*

The above three networks depict Example 1. The network in (A) is pairwise (Nash) stable for some parameters and the network in (B) is the only kind of feasible deviation. The network in (C) is an efficient network.

two agents forming a link across types and both participating agents delete a link. Such a move could be agents 1,6 forming a link while deleting their respecetive links to agents 3 and 4. We denote this network $\hat{\mu} = \mu \cup \{16\} \setminus \{13, 46\}$ *and we plot it in Figure 1.B. Benefits for agents 1 and 6 from network* μ *and deviating from it are:*

$$\begin{aligned} u_1(\hat{\mu}) + u_6(\hat{\mu}) &= (1+\delta) \cdot [z(\bar{x}, \bar{x}) + z(\underline{x}, \underline{x})] + [1 + 2\delta] \cdot [z(\bar{x}, \underline{x}) + z(\underline{x}, \bar{x})], \\ &= (1+\delta) \cdot \tfrac{1}{2} \cdot [Z(\bar{x}, \bar{x}) + Z(\underline{x}, \underline{x})] + [1 + 2\delta] \cdot Z(\bar{x}, \underline{x}), \\ u_1(\mu) + u_6(\mu) &= 2 \cdot [z(\bar{x}, \bar{x}) + z(\underline{x}, \underline{x})] = Z(\bar{x}, \bar{x}) + Z(\underline{x}, \underline{x}). \end{aligned}$$

We can express the condition that the deviation to $\hat{\mu}$ *is not pairwise profitable as:* $u_1(\mu) + u_6(\mu) > u_1(\hat{\mu}) + u_6(\hat{\mu})$. *This condition is sufficient for pairwise stability due to payoff symmetry in* μ *and no transfers.*

We now turn to deriving the condition for segregation to be inefficient. The aggregate benefits over all agents of the two networks, μ *and* $\tilde{\mu}$, *is expressed below in the two equations.*

$$\begin{aligned} U(\tilde{\mu}) &= (2+\delta) \cdot [Z(\bar{x}, \bar{x}) + Z(\underline{x}, \underline{x})] + [2 + 7\delta] \cdot Z(\bar{x}, \underline{x}), \\ U(\mu) &= 3 \cdot [Z(\bar{x}, \bar{x}) + Z(\underline{x}, \underline{x})]. \end{aligned}$$

---

link is positive it follows that a move consisting only in deletion of a link always leads to a loss. Thus, only coalition moves where new links are formed can be valuable. All links to same type agents are already formed in network μ. Therefore, the only feasible move consists in forming a link to agents of the other type.

Sorting is inefficient when: $U(\mu) < U(\tilde{\mu})$*. The two inequalities governing pairwise stability and inefficiency have the following positive solution:*

$$\begin{aligned} \delta^{stab}(\hat{Z}) &= \frac{\hat{Z}}{\hat{Z}+1}, \\ \delta^{opt}(\hat{Z}) &= \frac{\hat{Z}}{\hat{Z}+\frac{9}{2}}, \end{aligned} \qquad \hat{Z} = \frac{Z(\bar{x},\bar{x})+Z(\underline{x},\underline{x})}{2Z(\bar{x},\underline{x})} - 1,$$

where δ^{opt} *and* δ^{stab} *are thresholds for, respectively, when network* μ *becomes inefficient, and unstable when* δ *increases.*

The example above demonstrates that sorting can be inefficient when there are network effects despite there being complementarity in type, i.e. supermodular link values. The inefficiency stems from a novel source - the pairwise formation of links. The intuition is that under pairwise deviation the two agents do not internalize the total value created for the other agents; number of indirect links between a high and a low agent. Note that the above example has a close correspondence to Propositions 1 and 6 from de Martí and Zenou (2017) and that their results also holds only for cliques with very few agents (≤ 5), see literature review.

We proceed with a generalization of the example above which holds for various structures of the subnetworks within types and for multiple types. The aim is to extend the above example to a less restrictive setting for sorted connectivity. Below is our first general result where we establish sufficient and necessary conditions for the existence of networks with sorted connectivity.

Proposition 3 (Existence)**.** *The set of sort-connected networks is non-empty if and only if all of the following conditions hold:*

i) there is self-sufficiency for each type;

ii) more than one partner is allowed.

iii) either the degree quota is even or there is an even number of agents of each type;

The conditions in Proposition 3 are listed in order of importance. The essential condition is self-sufficiency, which ensures that there are enough links for each type to perfectly sort. The second condition of the degree quota exceeding one, is obvious as otherwise the problem would reduce to a simple one-to-one matching problem and type connectivity would not be possible. The final condition requiring either even numbered quota of links or even number of agents for each type is a little subtle. The reason is technical; if both of these conditions are not met then the total demand for links of the same type is uneven when there is no link surplus but each link takes up a capacity of two and thus must be an even number; the implication is that perfect sorting and no link surplus is incompatible when this even number condition is not met. We discuss the choice of equilibrium concepts in the discussion found in Section 5.

We move on to investigate stability and optimality of the network structure. We now generalize the thresholds from Example 1. These thresholds for stability and optimality hold for any number of types and number of agents for each type. Note that for optimality the value provided below is an upper bound of threshold value.

$$\delta^{stab} = \min_{x,\tilde{x}\in X}\left(\frac{\hat{Z}_{x,\tilde{x}}}{\hat{Z}_{x,\tilde{x}}+\max(n_x,n_{\tilde{x}})-|n_x-n_{\tilde{x}}|\cdot\hat{z}_{x,\tilde{x}}}\right), \qquad \hat{z}_{x,\tilde{x}} = \frac{z(x,\tilde{x})}{Z(x,\tilde{x})} \tag{6}$$

$$\bar{\delta}^{opt} = \min_{x,\tilde{x}\in X}\left(\frac{\hat{Z}_{x,\tilde{x}}}{\hat{Z}_{x,\tilde{x}}+\frac{1}{2}n_x n_{\tilde{x}}}\right), \qquad \hat{Z}_{x,\tilde{x}} = \frac{Z(x,x)+Z(\tilde{x},\tilde{x})}{2Z(x,\tilde{x})} - 1 \tag{7}$$

Using the first threshold above we can express our main result on the stability of sort-connected networks. Note that an alternative version of the above theorem under constant decay can be found in Appendix Appendix B in Theorem 3.

Theorem 1 (Stability). *Suppose there is supermodularity, then every sort-connected network is pairwise stable if $\delta \leq \delta^{stab}$; moreover, if there are only two types and sort-connected networks exist, then every pairwise network is also sort-connected.*

We have shown general stability of sort-connected networks. Moreover, when there are only two types, our results from Theorem 1 are substantially stronger. We establish that sorting is the unique pairwise stable outcome for low to moderate levels of externalities (i.e. $\delta < \delta^{stab}$).

We emphasize that Theorem 1 and δ^{stab} have implications for understanding the instability of perfect sorting in networks. For sufficiently high levels of spillover, i.e. $\delta > \delta^{stab}$, it holds that sorted networks are never pairwise stable. The reason is that agents of different types can benefit jointly by mutually forming a link and each breaking a same type link.

One can view Theorem 1 as generalizing not only Example 1 but also Propositions 1.ii and 4.iii from de Martí and Zenou (2017), who require that there are very few agents, two types and all agents of a given type link with one another.

We move on to discussing another main property of sort-connected networks, namely optimality, i.e. whether the network structure is efficient. In order to state our results we introduce a related network which has efficiency properties for moderate to high strength of utility from indirect connections.

Definition 8. *Let a* ***bridged, sort-connected*** *network be a sort-connected network where (i) for at least two types exactly one link is broken, (ii) each agent with a broken link forms exactly one link to other agents across types who also have a link broken.*

It is important to understand that in our model moving to a bridged, sort-connected network require two links to be established across types from a sort-connected network. This is a technical condition stemming from the fact that reducing the number of links among the same type by one frees up the capacity to establish a link by two agents; as a consequence it is possible for two links across types to be established. Using both of these possible links is important for establishing efficiency. It will turn out to also be important in the investigation of policy, see Proposition 4.

Theorem 2 (Efficiency)**.** *Suppose there is supermodularity, then the thresholds for efficiency and stability satisfy* $\bar{\delta}^{opt} < \delta^{stab}$*. Any sort-connected network is inefficient when* $\delta < \delta^{opt}$ *where it holds that* $\delta^{opt} \leq \bar{\delta}^{opt}$*. Finally, if there are two types then, sort-connected network are efficient when* $\delta \leq \bar{\delta}^{opt}$*, while bridged, sort-connected networks are efficient for* $\delta \geq \bar{\delta}^{opt}$*.*

The above theorem generalizes Example 1 by showing that under-connected networks with too little linking across sorted groups of agents occur more generally. It also extends Proposition 6 from de Martí and Zenou (2017) by removing the restriction to two types and linking between all same type agents as well as doing away with the limitation of having very few agents. Again, with only two types of agents the results are considerably stronger. We can show that the threshold for inefficiency now governs whether it is the sort-connected network or the bridged, sort-connected network that is efficient.

A visualization of the computed thresholds of externalities when there are two types is found in Figure 2. The thresholds are computed for varying population size and varying strength of complementarity. These plots can be seen as providing comparative statics along these two dimensions. The upper part of the figure keeps the population size fixed while lower ones keep the complementarity strength fixed. From inspection it is evident that both of the connection thresholds are approximately linear in log-log scale. This pattern suggests that both of the thresholds follow power-laws in the number of agents and strength of complementarity. Note that it is also straightforward to mathematically derive these patterns from the threshold definitions. Note also that it is possible to do a comparative static in the number of types. For example, if we assume that all same type agents get a certain payoff and cross type relations get some fraction of that, as in Jackson and Rogers (2005), then increasing the number of islands keeps the gain from pairwise deviation constant, but the gains in efficiency increase.

The remainder of this subsection will sketch a policy intervention that can mitigate the problem of suboptimal sorting by improving welfare through encouraging connection. These interventions can be seen more generally as a design problem, where the policy maker intervenes to induce a network that produces higher welfare. The tool that the policy maker employs is providing incentives to agents for forming specific links. Note that two agents of each type may need to be compensated. This stems from the fact that when compensating one agent to establish a link across types the agent to which it has deleted a link has an incentive to form a new link, which will potentially destabilize the sub-networks for each type. We discuss this assumption and how it relates to our choice of model in Section 5.

Define a *link-contingent contract* as a non-negative transfer, $\mathcal{C}_{ij}$, paid to i for linking with another agent j. Denote the vector of link-contingent contract as $\mathcal{C}$. We start demonstrating our results on intervention through a continuation of Example 1.

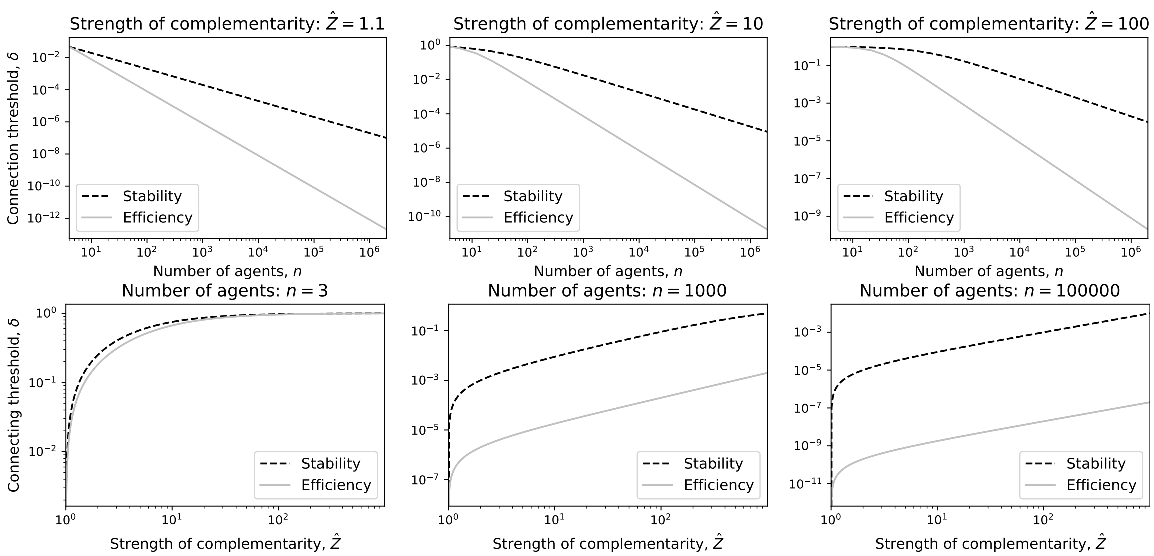


Figure 2: *Thresholds for connecting.*

Visualization of thresholds for connecting from Theorems 1 and 2. The upper part shows varying sizes of populations and fixed strength of complementarity. The lower part has varying strength of complementarity and fixed population sizes. It is assumed that there are two types which have an identical number of agents.

Example 1 (continued) *Let strength of utility from indirect connections satisfy* $\delta < \delta^{stab}$ *and assume that there is a policy maker who can issue link-contingent transfers as follows. We suppose that the policy maker offers conditional transfers such that the pairwise net-utility for agents 1,6 and agents 3,4 from forming a link satisfies:*

$$\mathcal{C}_{ij}+\mathcal{C}_{ji} = -\left[(1-\delta)\cdot(z(x,x)+z(\tilde{x},\tilde{x}))+(1+2\delta)\cdot z(x,\tilde{x})\right]+\varepsilon, \quad ij \in 16, 34, \quad \varepsilon > 0$$

Agents 1,6 and 3,4 have an incentive to form a link and break their existing links (i.e. $\{13,\ 46\}$*). This implies that a deviation from the sort-connected network* μ *to the bridged, sort-connected network* $\tilde{\mu}$ *is possible; see the networks depicted in Figure 1. We round off the example by noting that if* $\delta > \delta^{opt}$*, then the deviation to the bridged, sort-connected network raises aggregate utility.*

We now generalize the insight from the continuation of Example 1 into the following proposition. This proposition holds for an arbitrary number of agents when there are two types. We need to define an auxiliary term to describe the interventions:

Definition 9. *Let a network* $\tilde{\mu}$ *be* ***implementable*** *from* μ, τ *given* $\mathcal{C}$ *if there exist a sequence of tuples* $(\mu_0, \tau_0), .., (\mu_l, \tau_l)$ *where* $\mu_0 = \mu$*,* $\mu_l = \tilde{\mu}$*, and* $\tau_l = \tau$ *such that: for* $q = 1, .., l$ *from* μ_{q-1} *to* μ_q *is a feasible pairwise move which increases the pair's net-utility most given* $\mathcal{C}$*, and;* $\tilde{\mu}$ *is pairwise stable given* τ_q *and* $\mathcal{C}$*.*

Proposition 4 (Implementation)**.** *Suppose that there are two types, supermodularity and* $\delta < \delta^{stab}$*. It follows that a policy maker can implement a bridged, sort-connected network from any sort-connected network.*

The above result shows it is possible to have agents deviate to implement the bridged, sort-connected networks by offering link-contingent contracts. By combining the proposition with earlier results on inefficiency, it follows directly that:

Corollary 1. *If conditions for Proposition 4 hold and* $\delta > \bar{\delta}^{opt}$ *it follows that implementening the bridged, sort-connected network will result in higher welfare.*

The intuition of the corollary is that efficiency can be restored by compensating certain agents. Recall that in our model conditionality of transfers exist only between agents who are linked. This lack of conditionality for third parties implies that agents cannot fully internalize positive spillovers. Therefore, it is not surprising that efficiency is restored when allowing a third party, i.e. the policy maker, to transfer utility conditional on certain links as has been shown previously (Bloch and Jackson, 2007). We note that Proposition 4 and Corollary 1 outline a centralized intervention by a policy maker but it could also have been solved decentrally through conditional transfer by other agents, as in the framework of Bloch and Jackson (2007).

We note that the individual compensation paid to agents for connecting to others may not be equal. In particular, the payment may also depend on the types. This is the case when there is both supermodularity and monotonicity in Z. If these conditions hold, then agents may receive compensation that is increasing with their type.

4.2. Infinite population - constant decay in spillovers

We finalize this section by investigating what pattern of linking is exhibited when the count of agents becomes asymptotically infinite. In this large matching market we examine asymptotic perfect sorting, i.e. when the measured share of links to same-type agents converges to one. We employ the constant decay to measure utility from indirect connections as hyperbolic decay yield infinite payoff for any connected network with infinite number of agents. Note that we also use a stronger equilibrium concept, strong stability, which allows for coordination between coalitions of any size.

Definition 10. *Let* ***asymptotic perfect sorting*** *hold for a sequence of networks sets, M_n, if for any network, $\mu \in M_n$, where $n \to \infty$, it holds that $|\{ij \in \mu : x_i = x_j\}|/|\mu| \simeq 1$.*

Define *asymptotic independence* as $\delta < (\kappa - 1)^{-1}$. For large matching markets the sufficient conditions for asymptotic perfect sorting to emerge in strongly stable networks are:

Proposition 5. *If there is supermodularity, a degree quota and constant decay with asymptotic independence, then there is asymptotic perfect sorting for strongly stable networks.*

The result above demonstrates that the availability of many agents for linking induces perfect sorting in strongly stable networks. It demonstrates the same prediction as the conclusion of Becker (1973) for the marriage market model but holds in the presence of externalities with constant decay.

For deriving the result we exploit strong stability which implies that efficiency holds. Therefore, it is sufficient to show that asymptotic efficiency requires asymptotic perfect sorting. Although efficiency is a unique property for strong stability (and does not hold for weaker concepts) it can be argued that strong stability should be seen as a refinement with desirable properties which makes it more likely when it exists.[16]

We conclude this section by noting that we may interpret the result on sorting for infinite populations differently; there is no loss from sorting when there are many agents.

5. Concluding discussion

We have extended the assortative matching framework to a setting of networks. We have shown that in a general context that Becker's condition for sorting is still essential for stability. However, the same condition is insufficient for efficiency (when there is a finite population). The context is where types have enough members to form a community among themselves. We have sketched a policy that can help overcome this issue.

We have chosen to model costs implicitly via a degree quota in order to have comparability with the matching literature. We expect, however, that our results should easily translate to the standard connections model of Jackson and Wolinsky (1996). In this other setup we expect that the intuition should transfer when limiting the number of sub-networks within types to be either cliques or stars, as in de Martí and Zenou (2017). One advantage of translating the setup to the linear cost framework of the standard networks literature would be that the technical assumption of either even degree or an even number of agents for each type would not be necessary. Under hyperbolic decay one would also get a more natural efficient policy solution requiring only a single agent of each type to bridge the gap between their respective subnetworks.

Our analysis is based on other strict assumptions which we now review. We begin by noting that search frictions are important and have received attention in the literature (Chade et al., 2017) but for the sake of tractability we focus on a frictionless model. There are also a number of restrictive assumptions on payoff. The most crucial assumptions are payoff separability and fixed structure of externalities. Further research could explore how results generalize to less restrictive utility from indirect connections captured by the decay parameter, δ. For instance, does there exist a set of criteria that are more general than constant or hyperbolic decay for which our results hold. It is likely that our results are robust to including utility from

[16] In some circumstances the existence of contracts where an agent may subsidize or penalize another agent's link formation with alternative agents may imply strong stability even if contracts were limited to being pairwise specified, cf. Bloch and Jackson (2007).

network measures, e.g. triadic closure/clustering, that are common in the literature within economics on networks. Other critical assumptions are supermodularity and perfect transferability. Nevertheless, as mentioned in the introduction, these two assumptions can be replaced by monotonicity in individual link values and perfect non-transferability, which is also in line with some research on peer effects.[17] Finally, the model relies on some agents being of different types but it should be possible to remove this assumption.[18]

Another caveat with our analysis, and stable networks in general, is that these networks may not exist. We have shown some properties of existence under regularity conditions of sort-connected networks, see Propositions 3 and 1. However, beyond sort-connected networks we do not offer much in the case of externalities. The gross substitutes conditions from Kelso and Crawford (1982), which ensure existence of stable matchings in related settings, are not satisfied in our setting with externalities.[19] Nevertheless, by changing the equilibrium concept we expect that some of the lack of existence could be solved. One approach is using farsighted stability, as in Chwe (1994); Dutta et al. (2005); Herings et al. (2009). Another approach is using some approximative equilibrium concept e.g. cost of stability (the necessary payments to induce stability) from Bachrach et al. (2009).

Appendix A. Proofs

Lemma 1. *In the absence of network externalities then the set of strongly stable networks is equivalent to the set of pairwise stable networks.*

Proof. By definition it holds that any strongly stable network is pairwise stable. Thus, we need to show that any pairwise network is strongly stable. This claim is shown using similar to arguments to Klaus and Walzl (2009)'s Theorem 3.i.

Let μ with associated contracts τ be a network which is blocked by a coalition. It will be shown that for every coalition $t \in T$ that blocks, within the coalition there is a subset of no more than two members that also wishes to block the network. Let $\tilde{\mu}$ be the alternative network that the blocking coalitions implements through a feasible coalition move and τ be the transfers associated with $\tilde{\mu}$.

[17] Or, more broadly, by generalized increasing in differences from Legros and Newman (2007).

[18] The more extreme case is where all agents have different types, e.g., they exist in a ring with local complementarities, similar to Johnson and Gilles (2000) who assume agents' types are defined on a line. In this setting it may be that stable networks have the property that agents only link with the most similar agents and thus fail to connect with those further away.

[19] The lack of gross substitutes is due to the fact that a change in one active link can imply a change in the value of other links. This fact will violate gross substitutes.

It is always possible to partition the set of deleted links $\mu \backslash \tilde{\mu}$ into two: (i) a subset denoted $\hat{\mu}$ where for each link ij that can be deleted where one of the two partners can benefit, i.e. it holds that either $z_{ij} + \tau_{ij} - [c_i(\mu) - c_i(\mu \backslash ij)] < 0$ or $z_{ji} + \tau_{ji} - [c_j(\mu) - c_i(\mu \backslash ij)] < 0$; (ii) a subset denoted $\breve{\mu}$ where for each link ij neither of the previous two inequalities are satisfied.

Suppose that the first partition is non-empty, i.e. $\hat{\mu} \neq \emptyset$. However, as deleting links can be done by a single agent on its own then the move only takes needs the coalition of that agent to delete the link. Thus any part of a coalitional move that only involves profitably removing links can be performed in parts by a coalition with a single agent - therefore this move is also a pairwise block.

Thus it remains to be shown that the remaining part of coalitional move also can be performed as a pairwise block, i.e. when forming $\tilde{\mu} \backslash \mu$ and deleting $\breve{\mu}$. This part of the coalitional move must entail forming links as no links can be deleted profitably. The set of formed links $\tilde{\mu} \backslash \mu$ can be partitioned into a number of $|\tilde{\mu} \backslash \mu|$ feasible submoves of adding a single link while deleting links by each of the agents i and j who form a link. The feasibility for each of the partitioned moves is always true when there is a cost function as moves are unrestricted. It is now argued that each of the partitioned moves are feasible when there is a degree quota. If the network $\mu \cup ij$ is feasible then the move of simply adding the link is feasible. If $\mu \cup ij$ is not feasible, then agents i and j can delete at most one link each and if both μ and $\tilde{\mu}$ are feasible then this also feasible as the degree quota is kept.

For the coalitional move to $\tilde{\mu}$ it must be that at least at least one link among the implemented links $\tilde{\mu} \backslash \mu$ has a strictly positive value that exceeds the loss from deleting at most one link for each of two agents forming the link. This follows as it is known that deleting one or more links cannot add any value and thus must have weakly negative value and that by definition the total value to the blocking coalition must be positive. As every one of the partitioned moves is feasible, it follows that for every coalitional move there are two agents who can form link while potentially destroying current links and both be better off. In other words, for every coalition that blocks, there is a pairwise coalition that blocks. □

Proof of Proposition 1.. Suppose the claim is false. Let q be the lowest index for which the condition fail: for all $l < q$ it holds that $\mathcal{X}(\nu_i(\mu)/\{j\})_l \geq \mathcal{X}(\nu_j(\mu)/\{i\})_{l+l^*}$ where $l^* = \max(k_j(\mu) - k_i(\mu), 0)$. Thus there are two agents i', j' such that:

$$\begin{aligned} x_{j'} &= \mathcal{X}(\nu_j(\mu))_q, \quad j' \in (\nu_j(\mu) \backslash (\nu_i(\mu) \cup \{i\})), \\ x_{i'} &< \mathcal{X}(\nu_j(\mu))_q, \quad i' \in (\nu_i(\mu) \backslash (\nu_j(\mu) \cup \{j\})). \end{aligned}$$

Recall $k^* = \min(k_i(\mu), k_j(\mu))$. The argument why there must exist an agent i' in $\nu_j(\mu)$ but not in $(\nu_j(\mu) \cup \{j\})$ is that $|\{\iota \in \nu_i(\mu) : x_\iota < x_{j'}\}| > |\{\iota \in \nu_j(\mu) : x_\iota < x_{j'}\}|$.

This follows as by construction it holds that $|\{\iota \in \nu_i(\mu) : x_\iota < x_{j'}\}| = k^* - q + 1$ and $|\{\iota \in \nu_j(\mu) : x_\iota < x_{j'}\}| \leq k^* - q$.

The agents are such that $x_i > x_j, x_{i'} < x_{j'}$ as well as $ij', ji' \notin \mu$. However, this fact implies that there is a violation of strong stability: agents i, i', j, j' can deviate by destroying $\{ij, i'j'\}$ and forming $\{ij', i'j\}$ and thus increase payoffs due to supermodularity (cf. Equation 2). From Lemma 1 it follows that pairwise stability is also violated if strong stability is violated. ■

Proof of Proposition 2:. Condition (i) follows from the fact that it is possible for every agent to be linked with one another. Moreover every link adds value. Thus as a consequence every link can be formed and will add value both for the pair forming and it at the aggregate level; thus the unique pairwise and efficient outcome must be the complete network.

We move on to proving condition (ii). Suppose μ is a network where every agent has $n_{\bar{x}} - 1$ same-type links and $\kappa - n_{\bar{x}} + 1$ cross-type links.

Efficiency of μ follows from three facts. Firstly, μ the maximum distance of 2 between any two agents as all same-type links are active and all agents have at least one cross-type link; thus the potential benefits from indirect connections are maximized (both for constant and hyperbolic decay). Secondly, the number of same type links are maximized for all agents and this will maximize the benefits from direct links; thus there must exactly $n_{\bar{x}} - 1$ same type links. Finally, there can be no link surplus because violation there exist a network where every agent has $n_{\bar{x}} - 1$ same-type links and $\kappa - n_{\bar{x}} + 1$ cross-type links and thus has no link surplus; this must have strictly higher aggregate utility as every direct link increases utility.

Stability of μ follows from reviewing the feasible deviations. Let there be no transfers between any agents. Firstly, deleting one or more links is profitable as it lowers the agents own welfare. Secondly, forming a link requires deletion of one or more links by both agents. Deleting more links than one will lower the utility this only the deviations with deletion of a single link are relevant to consider - this corresponds to substitution of a link. Substituting either a same type link for another same type link or a cross type link for another cross type provides no change of utility to the pair of agents deviating. Substituting a cross type link for a same type link is not feasible. Substituting a same type link for a cross type link will lower the utility as the indirect benefits are unchanged but the direct benefits must be lower on aggregate due to supermodularity. ■

Lemma 2. *For every κ, n such that $n > \kappa$ and $n \cdot \kappa$ is even there exists a network $\mu_{n,\kappa}$ where all agents have exactly κ neighbors. Moreover, if $\kappa \geq 2$ then $\mu_{n,\kappa}$ is connected.*

Proof. Suppose n is even. Let % be the modulus operator. We can construct the following networks.

$$\hat{\mu}_{n,\kappa} = \left\{ij : i \in \{1, ..., \tfrac{n}{2}\},\, j \in \{(\tfrac{n}{2} + i \,\%\, \tfrac{n}{2}), ..., (\tfrac{n}{2} + [i + \kappa - 1] \,\%\, \tfrac{n}{2}\})\right\},\ \kappa \leq \tfrac{n}{2},$$

$$\tilde{\mu}_{n,\kappa} = \begin{cases} \hat{\mu}_{n,\kappa}, & \kappa \leq \frac{n}{2}, \\ \mu_c \backslash \hat{\mu}_{n,n-\kappa-1}, & \kappa > \frac{n}{2}. \end{cases}$$

Letting $\mu_{n,\kappa} = \tilde{\mu}_{n,\kappa}$ is sufficient for n is even. When n is odd we know that κ is even and thus we can use the following amended procedure instead:

$$\iota_{n,\kappa}(\iota) = \begin{cases} \frac{n-1}{2} + \iota, & \kappa \leq \frac{n-1}{2} \\ \frac{n-1}{2} + (\iota + \kappa) \,\%\, \frac{n-1}{2}, & \kappa > \frac{n-1}{2} \end{cases}$$

$$\mu_{n,\kappa} = \tilde{\mu}_{n-1,\kappa} \backslash \left\{ij : i \in \{1, .., \tfrac{\kappa}{2}\},\, j = \iota_{n,\kappa}(i)\right\} \cup \left\{ij : i = n,\, j \in (\cup_{\iota\{1,..,\frac{\kappa}{2}\}} \{\iota, \iota_{n,\kappa}(\iota)\})\right\}$$

We now show that if $\kappa \geq 2$ it follows that $\tilde{\mu}_{n,\kappa}$ is connected. Assume that n is even and suppose $\kappa \leq \frac{n}{2}$; for any $i \in N : i < \frac{n}{2}$ where $i' = i + 1$ and let $j = \frac{n}{2} + i + 1$ where $ij, i'j \in \tilde{\mu}_{n,\kappa}$; thus for all $i, i' \in \{1, .., \frac{n}{2}\}$ it holds that $p_{ii'}(\tilde{\mu}_{n,\kappa}) < \infty$. In addition, as for any $i \in N : i \leq \frac{n}{2}, j = \frac{n}{2} + i$ it holds that $ij \in \tilde{\mu}_{n,\kappa}$ it follows that $\tilde{\mu}_{n,\kappa}$ is connected. If instead $\kappa > \frac{n}{2}$ then by construction $ii' \in \tilde{\mu}_{n,\kappa}$ if either $\max(i, i') \leq \frac{n}{2}$ or $\min(i, i') > \frac{n}{2}$ as $ii' \notin \hat{\mu}_{n,n-\kappa-1}$. Moreover, for $i \in N : i < \frac{n}{2}$ and $j = \frac{n}{2} + (i + \kappa) \,\%\, \frac{n}{2}$ it holds that $ij \notin \hat{\mu}_{n,n-\kappa-1}$; thus $ij \in \tilde{\mu}_{n,\kappa}$. Therefore $\tilde{\mu}_{n,\kappa}$ must be connected.

Assume instead that n is odd. By the above argument there are at least two connected subnetworks consisting of agents in $\cup_{\iota\{1,..,\frac{\kappa}{2}\}} \{\iota, \iota_{n,\kappa}(\iota)\}$ and agents who are connected through agent, n, i.e. $N \backslash (\cup_{\iota\{1,..,\frac{\kappa}{2}\}} \{\iota, \iota_{n,\kappa}(\iota)\})$. If $\kappa \leq \frac{n-1}{2}$ where $i = \frac{\kappa}{2}$, $i' = \frac{\kappa}{2} + 1$ and $j = \frac{n-1}{2} + \frac{\kappa}{2} + 1$ then $ij, i'j \in \tilde{\mu}_{n,\kappa}$ and thus $\tilde{\mu}_{n,\kappa}$ is connected. If $\kappa > \frac{n-1}{2}$ where $i = \frac{\kappa}{2}$, $i' = \frac{\kappa}{2} + 1$ and $j = \frac{n-1}{2} + (\iota + \kappa + 1) \,\%\, \frac{n-1}{2}$ then $ij, i'j \in \tilde{\mu}_{n,\kappa}$ and thus $\tilde{\mu}_{n,\kappa}$ is connected. □

Lemma 3. *Suppose that* $\min_{x \in X} n_x > \kappa$, $\kappa \geq 2$. *If* $\exists i \in N$ *such that:*
a) $|\{i' \in \nu_i(\mu) : x_{i'} = x_i\}| \leq n_x - 2$*;*
b) $\min_{i' \in N_x \backslash \nu_i(\mu)} k_{i'}(\mu) = \kappa$*, and;*
c) $\max_{i' \in N_x \backslash \nu_i(\mu)} |\{i'' \in \nu_{i'}(\mu) : x_{i''} \neq x\}| = 0$*;*
then $\exists i', i'' \in \mu$ *such that* $i', i'' \notin \nu_i(\mu)$ *and* $p_{i'i''}(\mu \backslash \{i'i''\}) < \infty$

Proof. Suppose that for $i \in N$ the conditions a)-c) are met but the lemma is not true. If $i' \in N_x$ and $ii' \notin \mu$ then there must exist some $i'' \in N_x$ such that $i'i'' \in \mu$ and

$i'' \notin \nu_i(\mu)$ due to conditions a)-c). If $p_{i'i''}(\mu \backslash \{i'i''\}) < \infty$ then the proof is terminated so we must assume $p_{i'i''}(\mu \backslash \{i'i''\}) = \infty$.

As $p_{i'i''}(\mu \backslash \{i'i''\}) = \infty$ then the network $\mu \backslash \{i'i''\}$ has two components, $\mu', \mu'' \subseteq \mu \backslash (\{i'i''\}$, where in each component μ' or μ'' there are at least $\kappa + 1$ agents of type x (as for any $\iota \in (\nu_{i'}(\mu) \cup \nu_{i''}(\mu))$ it holds that $x_\iota = x$).

Agent i can at most be connected to one of i', i'' in $\mu \backslash (\{i'i''\})$ as otherwise i', i'' would be connected in $\mu \backslash \{i'i''\}$. Denote the in subnetwork of $\{\mu', \mu''\}$ where i is part of as $\tilde{\mu}$ and define $\tilde{N} = \{\iota \in N_x \backslash \nu_i(\mu) : \exists \iota' \in N : \iota\iota' \in \tilde{\mu}\}$.

Let $\iota_0 \in \arg\max_{\iota \in i', i''} p_{\iota i}$ and iteratively $\iota_l \in \nu_{\iota_{l-1}}(\mu), l \in \mathbb{N}$. Moreover, there must be a unique path in $\mu \backslash \{i'i''\}$ between any two agents $\iota, \iota' \in \tilde{N}$ as otherwise $i\iota, i\iota' \notin \mu$ but $p_{\iota\iota'}(\tilde{\mu} \backslash \{i'i''\}) < \infty$; by changing the labels we could denote $i' = \iota$ and $i'' = \iota'$ and we would have shown the existence of the desired pair of agents.

The fact here is a unique path between any two agents in $\tilde{N}$ entails that at level l or below there are $\sum_{q=0}^{l} (\kappa - 1)^q$ agents; thus $n_x \geq \sum_{q=0}^{l} (\kappa - 1)^q$. Let l be the minimal q such that $\forall \in \iota \in \tilde{N} : p_{i\iota} \leq q$; as n_x is finite such a q must exist. In addition, as there is a unique path between agents in μ then any agent $\iota \in \tilde{N} : p_{\iota\iota_0} = l$ has only one link, and thus its degree is less than κ (as $\kappa \geq 2$). This violates the condition that all $i' \in N$ where $x_{i'} = x$ has $k_{i'} = \kappa$. □

Proof Proposition 3.. The sufficiency of the conditions follows from Lemma 2 which can be applied to the subset of agents associated with each type as $\forall x \in X : n_x > \kappa$ and $\kappa \cdot n_x \in 2\mathbb{N}$.

The necessity of the conditions are straighforward. If either condition i) or iii) are violated then perfect sorting is not consistent with no-link surplus. If condition ii) is violated then there can be no type connectivity. ■

Proof Theorem 1..

Networks with sorted connectivity are stable Suppose μ that has sorted connectivity. We will demonstrate there are thresholds on δ such that μ has pairwise stability. We're only interested in the minimal thresholds such that for all values of externalities below those then stability holds. Thus it is sufficient to evaluate the deviations from the network where the net gains are highest.

The losses from breaking a link $ij \in \mu$ can be shown to have bounded from below such that: $\geq \delta \cdot (1 - Z(x, x))$. Suppose that $n_x = \kappa + 1, x \in X$ then $\{ij \in \mu : x_i = x, x_{i'} = x\}$ is a clique (i.e. any i, i' of type x are linked). This entails that $p_{ii'}(\mu \backslash \{ii'\}) = 2$ and thus $p_{ii'}(\mu \backslash \{ii'\}) < \infty$. Suppose instead that $n_x > \kappa + 1, x \in X$ then by Lemma 3 there exists some i, i', both of type x such that $p_{ii'}(\mu \backslash \{ii'\}) < \infty$.

Thus when evaluating losses at the threshold we can assume that when deleting some link ij that i, j are connected in $\mu\backslash\{ij\}$. Although the length of the shortest paths may increase, there will still be an indirect connection and therefore no loss of utility for anyone but the two agents who lose their link. Therefore we assume throughout that when evaluating thresholds if ii' is deleted in a sort-connected network then only agents i, i', who must be of same type, will each lose $(1-\delta)\cdot z(x,x)$ while no other agents incur a loss.

Suppose two agents i, j of distinct types respectively $x, \tilde{x}$ deviate by forming a link and delete a link each from μ. The total loss for i and j for deleting a link each is:

$$(1-\delta)\cdot[z(x,x)+z(\tilde{x},\tilde{x})] = (1-\delta)\cdot(\hat{Z}_{x,\tilde{x}}+1)\cdot Z(x,\tilde{x}).$$

The benefit gained for agent i for establishing a link to j is $[1+(n_{\tilde{x}}-1)\cdot(1-\delta)]\cdot z(x,\tilde{x})$. Thus the total benefits gained for i and j from pairwise deviation can be bounded as follows:

$$\begin{aligned} & [1+(n_x-1)\cdot\delta]\cdot z(x,\tilde{x}) + [1+(n_{\tilde{x}}-1)\cdot\delta]\cdot z(\tilde{x},x), \\ = \ & \langle 1+[\max(n_x,n_{\tilde{x}}) - |n_x-n_{\tilde{x}}|\cdot\hat{z}_{x,\tilde{x}}-1]\cdot\delta\rangle\cdot Z(x,\tilde{x}). \end{aligned}$$

where $\hat{z}_{x,\tilde{x}} = \frac{z(\arg\min_{x,\tilde{x}} n_x, \arg\max_{x,\tilde{x}} n_x)}{Z(x,\tilde{x})}$.

We can derive the threshold for pairwise stability, see definition of $\hat{Z}$ from Eq. (7). :

$$\begin{aligned} (1-\delta)\cdot(\hat{Z}_{x,\tilde{x}}+1)\cdot Z(x,\tilde{x}) &= \langle 1+[\max(n_x,n_{\tilde{x}})-|n_x-n_{\tilde{x}}|\cdot\hat{z}_{x,\tilde{x}}-1]\cdot\delta\rangle\cdot Z(x,\tilde{x}), \\ (1-\delta)\cdot(\hat{Z}_{x,\tilde{x}}+1) &= \langle 1+[\max(n_x,n_{\tilde{x}})-|n_x-n_{\tilde{x}}|\cdot\hat{z}_{x,\tilde{x}}-1]\cdot\delta\rangle, \\ \hat{Z}_{x,\tilde{x}} &= \left[\max(n_x,n_{\tilde{x}})-|n_x-n_{\tilde{x}}|\cdot\hat{z}_{x,\tilde{x}}+\hat{Z}_{x,\tilde{x}}\right]\cdot\delta, \\ \delta &= \frac{\hat{Z}_{x,\tilde{x}}}{\max(n_x,n_{\tilde{x}})-|n_x-n_{\tilde{x}}|\cdot\hat{z}_{x,\tilde{x}}+\hat{Z}_{x,\tilde{x}}}. \end{aligned} \tag{A.1}$$

Thus we can establish a lower bound for δ^{stab} (i.e. the upper bound in δ for pairwise stability of μ) by taking the minimum of right-hand-side in Equation A.1; thus it follows that: $\delta^{stab} = \min_{x,\tilde{x}\in X}\left(\frac{\hat{Z}_{x,\tilde{x}}}{\max(n_x,n_{\tilde{x}})-|n_x-n_{\tilde{x}}|\cdot\hat{z}_{x,\tilde{x}}+\hat{Z}_{x,\tilde{x}}}\right)$. ■

Pairwise stable networks have sorted connectivity when there are two types We need to show that every pairwise stable network is sort-connected. As there are only two types it holds that $X = \{\underline{x}, \overline{x}\}$. The outline of the proof is the we show the conditions in the following order; we begin with perfect sorting, then no link surplus, and finally type connectedness.

Perfect sorting We begin by supposing that μ is not perfectly sorted. We will construct a sequence of feasible deviations and show that they are profitable. As this part of the proof has considerable length it will be split into multiple sub-parts with a label that makes it easier to navigate.

Sequence of deviations The sequence of feasible deviations will consist of splitting up links between agents of different type and matching at even steps agents of type x and at odd steps agents of type $\tilde{x}$.

We first define sequences of agents and of deviations as steps $q = 1, 2, .., l$ where l is the number of steps. At each step we define the types as $x_q = x, \tilde{x}_q = \tilde{x}$ if q is even else vice versa.

Let the sequence of agent pairs, $i_0 j_0, i_1 j_1, ...$ be defined as follows. Let agents $i_0, j_0 \in N$ be such that $x_i \neq x_j$ and $ij \in \mu$; such i_0, j_0 must exist if μ is not perfectly sorted. Without loss of generality let $x_{i_0} = x$ and $x_{j_0} = \tilde{x}$ where $x, \tilde{x} \in X$. At step $q \in \mathbb{N}$ let $\iota_q = i_{q-1}$ if q is even else denote $\iota_q = j_{q-1}$. Also let $\eta_q \in \{i_{q-1}, j_{q-1}\} : \eta_q \neq \iota_q$. The advantage of this notation it is easier to define which links are formed between same type agents. Note that by construction we have that $\iota_1 = i_0$ and $\eta_1 = j_0$ as well as $x_{\iota_q} = x_q$ and $x_{\eta_q} = x_q$.

Using the sequence of agent pairs we construct the sequence of deviations as follows.

- At every step $q = 1, .., l$ a link is $\iota_q \eta_q$ $(=i_q j_q)$ is broken. We assume that broken links are elements of the original set, i.e. $\iota_q \eta_q \in \mu$ and can only be broken once $\iota_q \eta_q \notin \cup_{m=1}^{q-1} \{\iota_m \eta_m\}$.

- At every step $q = 1, .., l$ a link is formed $\iota_q \iota'_q \notin \mu$. This corresponds to $i_q i_{q-1} \notin \mu$ if q is odd and $j_q j_{q-1} \notin \mu$ if q is even. We assume that formed links are not part of the original set μ and can only be formed once $\iota_q \iota'_q \notin \cup_{m=1}^{q-1} \{\iota_m \iota'_m\}$..

- Combining the broken and formed links we get the coalitional move relative to μ:

$$\Delta\mu_q = \mu \cup \{\iota_q \iota'_q\} \backslash \{\iota_q \eta_q, \iota'_q \eta'_q\}, \quad q = 1, .., l-1 \tag{A.2}$$

$$\Delta\mu_l = \begin{cases} \mu \cup \{\iota_l \iota'_l\} \backslash \{\iota_l \eta_l, \iota'_l \eta'_l\}, \; k_{\iota_l}(\mu) = \kappa \\ \mu \cup \{\iota_l \iota'_l\} \backslash \{\iota_l \eta_l\}, \; k_{\iota'_l}(\mu) < \kappa \end{cases} \tag{A.3}$$

We note that the above sequence exists as we can always pick $l = 1$ and the assumption of having a broken link and link formed are satisfied by assumption.

Feasible partners We have defined the sequences of agents and deviations. We now restrict the set of partners at each step q for agent ι_q for $q = 1, .., l$:

$$\begin{aligned} N_q &= \{\iota \in N : x_\iota = x_q\} \\ \hat{N}_q &= \{\iota \in N_q \backslash \{\iota_q\} : \iota\iota_q \notin \mu\} \end{aligned}$$

A property of $\hat{N}_q$ is that $\hat{N}_q \neq \emptyset$; this follows as $\min_{\hat{x} \in X} n_{\hat{x}} \geq \kappa + 1$. We will now show that our restrictions on partner set has implications at each step $q = 1, .., l$:

- Let $\iota'_q \in \hat{N}_q$; this implies that Eq. (A.4) holds. This follows as a violation of Eq. (A.4) would imply that some agent ι''_q of type x_q would connected only through η_q; thus ι_q could link with ι''_q instead of ι'_q and thus ι_q can keep all its connections to agents who it was already connected to via η_q.

$$|\{\iota \in N : x_\iota = x_q \wedge p_{\iota\iota_q}(\mu) < \infty \wedge p_{\iota\iota_q}(\mu \cup \{\iota_q\iota'_q\} \backslash \{\eta_q\iota_q\}) = \infty\}| = 0 \quad \text{(A.4)}$$

- Suppose that Equation A.5 is violated for for $q \in \{1, .., l\}$. This is equivalent to it holds for any $\iota'_q \in \hat{N}_q$ where $\eta'_q \in \nu_{\iota'_q}(\mu)$ that there is some other $\iota''_q \in \hat{N}_q$ such that $p_{\iota'_q\iota''_q}(\Delta\mu_q) = \infty$. Let $\iota_q^{(1)} = \iota'_q$. As Equation A.5 must hold for any $\iota'_q \in \hat{N}_q$ we can reproduce the argument iteratively and thus for $\iota_q^{(m)} \in \hat{N}_q, q \in \mathbb{N}$ there is some $\eta_q^{(m)} \in \nu_{\iota_q^{(m)}}(\mu)$ such that for some $\iota_q^{(m+1)} \in \hat{N}_q \backslash \{\iota_q^{(1)}, .., \iota_q^{(m)}\}$ it holds that $p_{\iota_q^{(1)}\iota_q^{(m+1)}}(\Delta\mu_q) = \infty$. However, as $n < \infty$ it follows that there for some $q \in \mathbb{N}$ that $\hat{N}_q \backslash \{\iota_q^{(1)}, .., \iota_q^{(m)}\} = \emptyset$. Thus let instead $\iota'_q = \iota_q^{(m)}$; for any $\eta'_q \in \nu_{\iota'_q}(\mu)$ there is no $\iota''_q \in \hat{N}_q$ such that $p_{\iota'_q\iota''_q}(\mu) = \infty$. This contradicts that Equation A.5 is violated for agent $\iota'_q = \iota_q^{(m)}$.

$$|\{\iota \in N : x_\iota = x_q \wedge p_{\iota\iota'_q}(\mu) < \infty \wedge p_{\iota\iota'_q}(\Delta\mu_q) = \infty\}| = 0 \quad \text{(A.5)}$$

- Suppose Eq. (A.4) and (A.5) hold. We can demonstrate a variation of Eq. (A.4) where $p_{\iota\iota_q}(\Delta\mu_q) < \infty$, i.e. the ι and ι_q are connected despite the deletion of $\eta'_q\iota'_q$, see Eq. (A.6) below. The argument why Eq. (A.6) holds is as follows. Suppose $\exists \iota \in N : p_{\iota\iota_q}(\mu \cup \{\iota_q\iota'_q\} \backslash \{\iota_q\eta_q\}) < \infty$ and $p_{\iota\iota_q}(\Delta\mu_q) = \infty$ and $x_\iota = x_q$. If $p_{\iota_q\iota'_q}(\mu) < \infty$ then as it also holds that $p_{\iota_q\iota'_q}(\Delta\mu_q) < \infty$ it follows that

$p_{\iota\iota'_q}(\mu) < \infty$ and $p_{\iota\iota'_q}(\Delta\mu_q) = \infty$ which violates Eq. (A.5). Thus it must be that $p_{\iota_q\iota'_q}(\mu) = \infty$. Suppose instead $p_{\iota_q\iota'_q}(\mu) = \infty$. Then it must be that $p_{\iota\eta'_q}(\mu) < \infty$ and thus $p_{\iota\iota'_q}(\mu) < \infty$ as $\iota'_q\eta'_q \in \mu$ which violates that $p_{\iota_q\iota''_q}(\mu) = \infty$.

$$|\{\iota \in N : x_\iota = x_q \wedge p_{\iota\iota_q}(\mu) < \infty \wedge p_{\iota\iota_q}(\Delta\mu_q) = \infty\}| = 0 \qquad \text{(A.6)}$$

Gains from deviation We now move on to describing the gains to individuals from deviating. We assume initial transfers satisfy:

$$-\tau_{i_0j_0} > (1-\delta)\cdot z(\tilde{x},\tilde{x}) - [1 + (n_{\tilde{x}} - 1)\cdot\delta]\cdot z(\tilde{x}, x). \qquad \text{(A.7)}$$

The above inequality must hold for either type x or $\tilde{x}$ as we substitute labels for i, j as well as $x, \tilde{x}$ due to $\Upsilon > 0$.

By inserting i, j for ι, η we yield the following expression:

$$\sum_{q=l'}^{l-1} \Delta U_q = \sum_{q=l'}^{l-2} \Delta\hat{U}_q + u_{\iota'_{l-1}}(\Delta\mu_{l-1}) - u_{\iota'_{l-1}}(\mu) + u_{\iota_{l'}}(\Delta\mu_{l'}) - u_{\iota_{l'}}(\mu) \qquad \text{(A.8)}$$

Suppose that at every $q \in \mathbb{N} : q < l$ it holds that $\iota'_q \notin \nu_{\iota_q}(\mu)$, $x_{\iota'_l} = x_l$ and let $\eta'_l \in \nu_{\iota'_l}(\mu)$.) and let:

$$\begin{aligned} \Delta U_q &= u_{\iota_q}(\Delta\mu_q) - u_{\iota_q}(\mu) + u_{\iota'_q}(\Delta\mu_q) - u_{\iota'_q}(\mu) && \text{(A.9)} \\ \Delta\hat{U}_q &= u_{i_q}(\Delta\mu_{q+\mathbf{1}_{q:even}}) - u_{i_q}(\mu) + u_{j_q}(\Delta\mu_{q+\mathbf{1}_{q:odd}}) - u_{j_q}(\mu) && \text{(A.10)} \end{aligned}$$

We define the auxiliary term Υ below which is useful for bounding the gains from deviation. As $\delta < \delta^{stab}$ it follows from Equation A.1 that $\Upsilon > 0$.

$$\Upsilon = (1-\delta)\cdot[z(\bar{x},\bar{x}) + z(\underline{x},\underline{x})] - [1 + (n_{\underline{x}} - 1)\cdot\delta]\cdot z(\underline{x},\bar{x}) - [1 + (n_{\bar{x}} - 1)\cdot\delta]\cdot z(\bar{x},\underline{x}). \qquad \text{(A.11)}$$

Gains for ι_l: As Eq. (A.4) holds it follows that net gains for ι_q from deleting the link with η_q while forming a link together with ι'_q can be bounded: the upper bound of losses is when a connection is lost to all agents of type $\tilde{x}_q$: $[1 + (n_q - 1)\cdot\delta]\cdot z(x_q, x_q)$; the lower bound of gains is $(1-\delta)\cdot z(x_q, \tilde{x}_q)$ as the distance between $\iota_q\iota'_q$ is shortened to 1.

$$u_{\iota_q}(\mu \cup \{\iota_q\iota'_q\} \backslash \{\iota_q\eta_q\}) - u_{\iota_q}(\mu) \;\geq\; (1-\delta)\cdot z(x_q, x_q) - [1 + (n_q - 1)\cdot\delta]\cdot z(x_q, \tilde{x}_q) \qquad \text{(A.12)}$$

Analogue to the derivation of Ineq. (A.12) the net gains are bounded when Eqs. (A.5) and (A.6) are satisfied:

$$\min_{\iota\in\{\iota_q,\iota'_q\}} [u_\iota(\Delta\mu_q) - u_\iota(\mu)] \geq (1-\delta)\cdot z(x_q, x_q) - [1 + (n_q - 1)\cdot\delta]\cdot z(x_q, \tilde{x}_q), \qquad \text{(A.13)}$$

The first and foremost implication of Ineq. (A.13) and the fact that $x_{i_q} \neq x_{j_q}$ is that:

$$u_{i_q}(\Delta\mu_{q+\mathbf{1}_{odd}(q)}) - u_{i_q}(\mu) + u_{j_q}(\Delta\mu_{q+\mathbf{1}_{even}(q)}) - u_{j_q}(\mu) \geq \Upsilon. \tag{A.14}$$

Another implication of Ineq. (A.7) when combined with Ineq. (A.13) is that:

$$\begin{aligned} u_{\iota_1}(\Delta\mu_1) - u_{\iota_1}(\mu) - \tau_{\iota_1\eta_1} &\geq (1-\delta)\cdot z(x,x) - [1+(n_x-1)\cdot\delta]\cdot z(x,\tilde{x}) - \tau_{i_0 j_0} \\ u_{\iota_1}(\Delta\mu_1) - u_{\iota_1}(\mu) - \tau_{\iota_1\eta_1} &\geq \Upsilon \end{aligned} \tag{A.15}$$

Furthermore, we can restrict transfers as follows. In order for $\Delta\mu_q$ not to be a profitable pairwise deviation it must hold that:

$$\begin{aligned} u_{\iota_q}(\mu) + u_{\iota'_q}(\mu) + \tau_{\iota'_q\eta'_q} + \tau_{\iota_q\eta_q} &\geq u_{\iota_q}(\Delta\mu_q) + u_{\iota'_q}(\Delta\mu_q) \\ \tau_{\iota'_q\eta'_q} &\geq \Delta U_q + \tau_{\eta_q\iota_q} \end{aligned}$$

We can rewrite the above inequality using that $\iota'_{q-1} = \eta_q, \eta'_{q-1} = \iota_q$ and thus $\tau_{\iota'_{q-1}\eta'_{q-1}} = \tau_{\eta_q\iota_q}$. We also substitute in Equation A.9 and assume the above inequality holds for any $q < l$:

$$\begin{aligned} \tau_{\iota'_{l-1}\eta'_{l-1}} &\geq \Delta U_{l-1} + \tau_{\iota'_{l-2}\eta'_{l-2}} \\ \tau_{\iota'_{l-1}\eta'_{l-1}} &\geq \sum_{q=l'}^{l-1} \Delta U_q + \tau_{\iota'_{l'-1}\eta'_{l'-1}} \end{aligned} \tag{A.16}$$

As $\tau_{\eta_l\iota_l} = \tau_{\iota'_{l-1}\eta'_{l-1}}$ and $-\tau_{\iota_l\eta_l} = \tau_{\eta_l\iota_l}$ it follows that using Equation A.8:

$$\begin{aligned} -\tau_{\iota_l\eta_l} &\geq \sum_{q=1}^{l-1} \Delta U_q + \tau_{\iota'_{l'-1}\eta'_{l'-1}} \\ &= \sum_{q=1}^{l-2} \Delta\hat{U}_q + u_{\iota'_{l-1}}(\Delta\mu_{l-1}) - u_{\iota'_{l-1}}(\mu) + u_{\iota_1}(\Delta\mu_1) - u_{\iota_1}(\mu) + \tau_{\iota'_0\eta'_0} \\ &= \sum_{q=1}^{l-2} \Delta\hat{U}_q + u_{\eta_l}(\Delta\mu_{l-1}) - u_{\eta_l}(\mu) + u_{\iota_1}(\Delta\mu_1) - u_{\iota_1}(\mu) - \tau_{\iota_1\eta_1} \end{aligned} \tag{A.17}$$

Gains for partners of ι_l: link surplus We will now examine and find bounds on the benefits of deviating when we assume that $k_{\iota'_l}(\mu) < \kappa$. As Eq. (A.4) holds it follows that

$$u_{\iota'_l}(\mu \cup \{\iota_l\iota'_l\}\backslash\{\iota_l\eta_l\}) - u_{\iota'_l}(\mu) \geq (1-\delta)\cdot z(x_l, x_l),$$

and thus $u_{\iota_l'}(\mu \cup \{\iota_l \iota_l'\} \backslash \{\iota_l \eta_l\}) > 0$.

$$\begin{aligned}
& u_{\iota_l}(\Delta\mu_l) - u_{\iota_l}(\mu) - \tau_{\iota_l \eta_l}, \\
\geq \quad & u_{\iota_l}(\Delta\mu_l) - u_{\iota_l}(\mu) + u_{\eta_l}(\Delta\mu_{l-1}) - u_{\eta_l}(\mu) + \sum_{q=1}^{l-2} \Delta\hat{U}_q + u_{\iota_1}(\Delta\mu_1) - u_{\iota_1}(\mu) - \tau_{\iota_1 \eta_1}, \\
\geq \quad & \sum_{q=1}^{l-1} \Delta\hat{U}_q + u_{\iota_1}(\Delta\mu_1) - u_{\iota_1}(\mu) - \tau_{\iota_1 \eta_1}. \qquad \text{(A.18)}
\end{aligned}$$

We now apply Ineqs. (A.14) and (A.15) to the above expression which implies that the gains for ι_l from deviating are bounded below by $l \cdot \Upsilon$. As we have that $\Upsilon > 0$ it follows that:

$$u_{\iota_l}(\mu \cup \{\iota_l \iota_l'\} \backslash \{\iota_l \eta_l\}) - u_{\iota_l}(\mu) - \tau_{\iota_l \eta_l} > 0 \qquad \text{(A.19)}$$

Combining that both ι_l, ι_l' have incentive to deviate it follows their joint deviation is profitable which violates pairwise stability. Thus it must be that $k_{\iota_l'}(\mu) = \kappa$.

Gains for partners of ι_l: dropping same type partner with no loss of connectivity Suppose there exists $\iota_l' \in \hat{N}_l, \iota_l'' \in \hat{N}_l \backslash \{\iota_l'\}$ such that $\iota_l' \iota_l'' \in \mu$, $p_{\iota_l' \iota_l''}(\mu \backslash \{\iota_l \eta_l, \iota_l' \iota_l''\}) < \infty$ and $\tau_{\iota_l' \iota_l''} \leq 0$. This entails that $u_{\iota_l'}(\Delta\hat{\mu}_l) - u_{\iota_l'}(\mu) \geq 0$ where $\Delta\hat{\mu}_l = \mu \cup \{\iota_l \iota_l'\} \backslash \{\iota_l \eta_l, \iota_l' \iota_l''\}$. This follows from $u_{\iota_l'}(\Delta\hat{\mu}_l) - u_{\iota_l'}(\mu) = u_{\iota_l'}(\Delta\hat{\mu}_l) - u_{\iota_l'}(\mu \cap \Delta\hat{\mu}_l) - [u_{\iota_l'}(\mu \cap \Delta\hat{\mu}_l) - u_{\iota_l'}(\mu)]$ and $u_{\iota_l'}(\Delta\hat{\mu}_l) - u_{\iota_l'}(\mu \cap \Delta\hat{\mu}_l) \geq 1 - z(x,x)$ and $u_{\iota_l'}(\mu \cap \Delta\hat{\mu}_l) - u_{\iota_l'}(\mu) = 1 - z(x,x)$. As $\tau_{\iota_l' \iota_l''} \leq 0$ it follows that that utility for ι_l' is:

$$u_{\iota_l'}(\Delta\hat{\mu}_l) - u_{\iota_l'}(\mu) - \tau_{\iota_l' \iota_l''} \geq 0.$$

And utility for ι_l can bounded be as follows using Inequality A.12 for $u_{\iota_l}(\Delta\hat{\mu}_l) - u_{\iota_l}(\mu)$ as Equation A.4 holds :

$$\begin{aligned}
& u_{\iota_l}(\Delta\hat{\mu}_l) - u_{\iota_l}(\mu) - \tau_{\iota_l \eta_l} \\
= \quad & u_{\iota_l}(\Delta\hat{\mu}_l) - u_{\iota_l}(\mu) + \tau_{\eta_l \iota_l} \\
\geq \quad & \sum_{q=1}^{l-1} \Delta U_q + u_{\iota_l}(\Delta\hat{\mu}_l) - u_{\iota_l}(\mu) + \tau_{j_0 i_0} \\
= \quad & \sum_{q=1}^{l-2} \Delta\hat{U}_q + u_{\iota_l}(\Delta\hat{\mu}_l) - u_{\iota_l}(\mu) + u_{\eta_l}(\Delta\mu_{l-1}) - u_{\eta_l}(\mu) + u_{i_0}(\Delta\mu_1) - u_{i_0}(\mu) - \tau_{i_0 j_0} \\
\geq \quad & l \cdot \Upsilon \\
> \quad & 0 \qquad \text{(A.20)}
\end{aligned}$$

The above inequalities entails that ι_l, ι'_l can deviate profitably pairwise; this is a violation of pairwise stability and thus cannot be true. Thus there exists no $\iota'_l\iota''_l \in \mu$ such that $\iota'_l \in \hat{N}_l, \iota''_l \in \hat{N}_l \backslash \{\iota'_l\}$ as well as $p_{\iota'_l\iota''_l}(\mu \backslash \{\iota_l\eta_l, \iota'_l\iota''_l\}) < \infty$ and $\tau_{\iota'_l\iota''_l} \leq 0$.

Gains for partners of ι_l: only same type partners. Suppose that $\forall \iota'_l \in \hat{N}_l : \nexists \eta'_l \in \nu_{\iota'_l}(\mu \backslash \cup_{q=1}^{l} \{\iota_q\eta_q\}) : x_{\eta'_l} \neq x_l$. This entails that $\forall \iota'_l \in \hat{N}_l : \nexists \eta'_l \in \nu_{\iota'_l}(\mu) : x_{\eta'_l} \neq x_l$ as $k_{\iota'_l}(\mu \backslash \cup_{q=1}^{l} \{\iota_q\eta_q\}) = k_{\iota'_l}(\mu)$. By Lemma 3 it follows there exists $\iota'_l, \iota''_l \in \hat{N}_l \backslash \nu_i(\mu)$ such that $p_{\iota'_l\iota''_l}(\mu \backslash \{\iota'_l\iota''_l\}) < \infty$, $\iota'_l\iota''_l \in \mu$ and $\tau_{\iota'_l\iota''_l} \leq 0$ which by the arguments above cannot be true. Therefore there has to exist some $\iota'_l \in \hat{N}_l$ for which there is an agent $\eta'_l \in \nu_{\iota'_l}(\mu \backslash \cup_{q=1}^{l} \{\iota_q\eta_q\})$ where it holds that $x_{\eta'_l} \neq x_l$.

Gains for partners of ι_l: link is already broken. We shown above that there must exist some partner η'_q of different type than ι'_q such that $\iota'_q\eta'_q \in \mu$. However, there can only be a finite number of such links. Therefore, after a number of broken links there will be only be duplicate links left, i.e. $\iota'_l, \eta'_l \in (\mu \cap \cup_{q=1}^{l}\{\iota_q\eta_q\})$. That is for some $l' < l$ it holds that either $\iota_l, \eta_l = \iota_{l'}, \eta_{l'}$ if $l - l'$ is even or $\iota_l, \eta_l = \eta_{l'}, \iota_{l'}$ if $l - l'$ is odd.

If $l - l'$ is odd, then $\tau_{\iota'_{l'-1}\eta'_{l'-1}} = -\tau_{\iota'_{l-1}\eta'_{l-1}}$ and therefore we can reduce the Inequality A.16:

$$
\begin{aligned}
0 &\geq \sum_{q=l'}^{l-1}[u_{\iota_q}(\Delta\mu_q) - u_{\iota_q}(\mu) + u_{\iota'_q}(\Delta\mu_q) - u_{\iota'_q}(\mu)] + 2\tau_{\iota'_{l'-1}\eta'_{l'-1}} \\
&= \sum_{q=l'}^{l-2} \Delta\hat{U}_q + u_{\iota'_{l-1}}(\Delta\mu_{l-1}) - u_{\iota'_{l-1}}(\mu) + u_{\iota_{l'}}(\Delta\mu_{l'}) - u_{\iota_{l'}}(\mu) + 2\tau_{\iota'_{l'-1}\eta'_{l'-1}} \\
&= \sum_{q=l'}^{l-2} \Delta\hat{U}_q + 2 \cdot \left\langle u_{\eta'_{l'-1}}(\Delta\mu_{l'}) - u_{\eta'_{l'-1}}(\mu) + \tau_{\iota'_{l'-1}\eta'_{l'-1}} \right\rangle \\
&= \sum_{q=l'}^{l-2} \Delta\hat{U}_q + 2 \cdot \left\langle u_{\eta'_{l'-1}}(\Delta\mu_{l'}) - u_{\eta'_{l'-1}}(\mu) + \sum_{q=1}^{l'-1} \Delta U_q + \tau_{\iota'_0\eta'_0} \right\rangle \\
&= \sum_{q=l'}^{l-2} \Delta\hat{U}_q + 2 \cdot \sum_{q=1}^{l'-1} \Delta\hat{U}_q + 2 \cdot [u_{\iota_1}(\Delta\mu_1) - u_{\iota_1}(\mu) - \tau_{\iota_1\eta_1}] \\
&\geq (l + l') \cdot \Upsilon \\
&> 0,
\end{aligned}
$$

thus there must be a feasible pairwise deviation for some agent pair ι_l, ι'_l where

$q \in [[1, l]]$.

If $l - l'$ is even then $\tau_{\iota'_{l-1}\eta'_{l-1}} = \tau_{\iota'_{l'-1}\eta_{l'-1}}$; thus Inequality A.16 for no pairwise deviation becomes: $0 \geq \sum_{q=l'}^{l-1} \Delta U_q$. This can in turn be rewritten as follows:

$$0 \geq \sum_{q=l'}^{l-2} \Delta \hat{U}_q + u_{\iota'_{l-1}}(\Delta\mu_{l-1}) - u_{\iota'_{l-1}}(\mu) + u_{\iota_{l'}}(\Delta\mu_{l'}) - u_{\iota_{l'}}(\mu)$$

Using that $\iota_q = \eta'_{q-1}$ and $\eta'_{l'-1} = \eta'_{l-1}$ we get: $0 \geq \sum_{q=l'}^{l-1} \Delta \hat{U}_q$. Recall that for all $q \in \mathbb{N} : q < l$ it holds that $\Delta \hat{U}_q \geq \Upsilon$ where $\Upsilon > 0$. Thus there must be a feasible pairwise deviation.

We have now shown that the network μ has perfect sorting.

<u>*No link surplus*</u> Suppose that μ has link surplus. This would entail that $\exists i \in N : k_i(\mu) < \kappa$. As $n_x > \kappa$ it must be that $\exists i' \in N : x_{i'} = x_i, ii' \notin \mu$. Suppose that $k_{i'} < \kappa$ then $\sum_{\iota \in \{i,i'\}}[u_\iota(\mu \cup \{ii'\}) - u_\iota(\mu)] > 0$ and thus ii' can be formed profitably pairwise. Moreover, as $k_{i'}(\mu) = \kappa$ it follows that $\exists i'' \in \nu_{i'} : ii'' \notin \mu$, $x_{i''} = x_i$. By Lemma 3 it follows there exists $\iota, \iota' \in \tilde{N} \backslash \nu_i(\mu)$ such that $p_{\iota\iota'}(\mu \backslash \{\iota\iota'\}) < \infty$, $\iota\iota' \in \mu$ and $\tau_{\iota\iota'} \leq 0$. This entails that $u_\iota(\mu) - u_\iota(\mu \backslash \{\iota\iota'\}) + \tau_{\iota\iota'} \leq (1-\delta)z(x,x)$. Moreover, as $\sum_{j \in \{i,\iota\}}[u_j(\mu \cup \{i\iota\} \backslash \{\iota\iota'\}) - u_j(\mu \backslash \{\iota\iota'\})] \geq (1-\delta) \cdot Z(x,x)$ it holds that:

$$\sum_{j \in \{i,\iota\}} [u_j(\mu \cup \{i\iota\} \backslash \{\iota\iota'\}) - u_j(\mu)] - \tau_{\iota\iota'} \geq (1-\delta) \cdot z(x,x)$$

.

Thus i, ι can deviate profitably pairwise which contradicts pairwise Nash stability. Therefore it must be that that μ has no link-surplus.

<u>*Type connected*</u> Suppose that μ is not type connected. As μ we have established perfect sorting and no-link surplus there exist $i, i', j, j' \in N : x_i = x_{i'} = x_j = x_{j'}$ and $ij, i'j' \in \mu$ and $p_{ii'}(\mu) = \infty$. Without loss of generality we assume that $\tau_{ij}, \tau_{i'j'} \leq 0$ (otherwise we could simply switch identities some i's and j's). This entails:

$$\min_{\iota \in \{i,i'\}} [u_\iota(\mu \backslash \{ij, i'j'\}) - u_\iota(\mu)] + \tau_{ij} + \tau_{i'j'} \leq 2(1-\delta) \cdot z(x,x)$$

Also we have that:

$$\min_{\iota \in \{i,i'\}} [u_\iota(\mu \cup \{ii'\} \backslash \{ij, i'j'\}) - u_\iota(\mu \backslash \{ij, i'j'\})] \geq (\kappa+1) \cdot (1-\delta) \cdot z(x,x)$$

This entails that $\sum_{\iota \in \{i,i'\}}[u_\iota(\mu \cup \{ii'\} \backslash \{ij, i'j'\}) - u_\iota(\mu)] - \tau_{ij} - \tau_{i'j'} \geq \kappa \cdot (1-\delta) \cdot Z(x,x)$; thus i, i' can deviate profitably. Thus we have established each of the three properties are necessary for pairwise stability. ■

Proof of Theorem 2.

Inefficiency of sort-connected networks We aim to prove that there exists a threshold δ^{opt} such that if $\delta > \delta^{opt}$ then there exists a network which has higher aggregate utility than any network with sorted connectivity.

Suppose μ is sort-connected and let $\hat{\mu}$ be a bridged, sort-connected network such that only two distinct types $x, \tilde{x}$ have links across. Denote two agents of type x who link across types as i, i' and those of type $\tilde{x}$ who link as j, j'. This entails that $\hat{\mu} = \mu \cup \{ij, i'j'\} \backslash \{ii', jj'\}$. It follows that the loss in aggregate utility is captured by Eq. (A.21). The gain in aggregate utility follows Eq. (A.22). The aggregate net-gain in utility is captured by Eq. (A.23).

$$U(\mu \backslash \{ii', jj'\}) - U(\mu) = -(1-\delta)[Z(x,x) + Z(\tilde{x},\tilde{x})] \tag{A.21}$$

$$U(\hat{\mu}) - U(\mu \backslash \{ii', jj'\}) = [\delta \cdot (n_x n_{\tilde{x}} - 2) + 2] \cdot Z(x,\tilde{x}) \tag{A.22}$$

$$U(\hat{\mu}) - U(\mu) = -(1-\delta)[Z(x,x) + Z(\tilde{x},\tilde{x}) - 2Z(x,\tilde{x})] + \delta n_x n_{\tilde{x}} \cdot Z(x,\tilde{x}) \tag{A.23}$$

The derivative of Eq. (A.23) wrt. δ is $Z(x,x) + Z(\tilde{x},\tilde{x}) - 2Z(x,\tilde{x}) + n_x n_{\tilde{x}} \cdot Z(x,\tilde{x})$ Due to supermodularity it holds that $Z(x,x) + Z(\tilde{x},\tilde{x}) - 2Z(x,\tilde{x}) > 0$. Therefore, $U(\hat{\mu}) - U(\mu)$ is monotone increasing in δ. Moreover, $U(\hat{\mu}) - U(\mu) = n_x n_{\tilde{x}} Z(x,\tilde{x}) > 0$ when $\delta = 1$ and $(\hat{\mu}) - U(\mu) = -[Z(x,x) + Z(\tilde{x},\tilde{x}) - 2Z(x,\tilde{x})] < 0$ when $\delta = 0$. As both aggregate losses and gains are continuous in δ it follows by the intermediate value theorem that that there exist a threshold δ^{opt} such that if $\delta > \delta^{opt}$ then $U(\hat{\mu}) > U(\mu)$.

For any two types we can compute a threshold $\delta^{opt}_{x,\tilde{x}}$ where gains equal losses as below. We use definition of $\hat{Z}$ from Eq. (7).

$$\begin{aligned} (n_x \cdot n_{\tilde{x}}) \cdot Z(x,\tilde{x}) \cdot \delta &= 2(1-\delta) \cdot (\hat{Z}_{x,\tilde{x}}) \cdot Z(x,\tilde{x}), \\ (\hat{Z}_{x,\tilde{x}} + \tfrac{1}{2} n_x \cdot n_{\tilde{x}}) \cdot \delta &= \hat{Z}_{x,\tilde{x}}, \\ \delta &= \frac{\hat{Z}_{x,\tilde{x}}}{\hat{Z}_{x,\tilde{x}} + \frac{1}{2} n_x \cdot n_{\tilde{x}}}. \end{aligned} \tag{A.24}$$

For each pair of types we can compare with the threshold for stability $\delta^{stab}_{x,\tilde{x}}$ from Eq. (A.1).

$$
\begin{aligned}
\delta_{x,\tilde{x}}^{stab} &> \delta_{x,\tilde{x}}^{opt} \\
\left(\frac{\hat{Z}_{x,\tilde{x}}}{\hat{Z}_{x,\tilde{x}}+\max(n_x,n_{\tilde{x}})-|n_x-n_{\tilde{x}}|\cdot\hat{z}_{x,\tilde{x}}}\right) &> \frac{\hat{Z}_{x,\tilde{x}}}{\hat{Z}_{x,\tilde{x}}+\frac{1}{2}n_x\cdot n_{\tilde{x}}} \\
\frac{1}{2}n_x\cdot n_{\tilde{x}} &> \max(n_x,n_{\tilde{x}})-|n_x-n_{\tilde{x}}|\cdot\hat{z}_{x,\tilde{x}} \\
\frac{1}{2}n_x\cdot n_{\tilde{x}}+|n_x-n_{\tilde{x}}|\cdot\hat{z}_{x,\tilde{x}} &> \max(n_x,n_{\tilde{x}})
\end{aligned}
$$

As it holds both that $\hat{z}_{x,\tilde{x}} > 0$ and that $n_x \cdot n_{\tilde{x}} > \max n_x, n_{\tilde{x}}$ (because $\min n_x, n_{\tilde{x}} >= 2$) it follows that $\delta_{x,\tilde{x}}^{opt} < \delta_{x,\tilde{x}}^{stab}$. As a consequence it must be that $\min_{x\neq\tilde{x}} \delta_{x,\tilde{x}}^{opt} < \min_{x\neq\tilde{x}} \delta_{x,\tilde{x}}^{stab}$. In other words, this implies that the dominance hold globally for the threshold $\bar{\delta}^{opt} < \delta^{stab}$. As we only evaluated bridged, sort-connected networks where two types link across there may exist lower thresholds for optimality $\delta^{opt} \leq \bar{\delta}^{opt}$. By construction it holds that $\delta^{opt} < \delta^{stab}$.

Efficiency of networks Our next aim is to show the following properties when there are only two types: (i) $\delta^{opt} = \bar{\delta}^{opt}$; (ii) for $\delta \leq \delta^{opt}$ it holds that any sort-connected network is efficient; and (iii) for $\delta \geq \delta^{opt}$ any bridged, sort-connected network is efficient.

Property (i). As there are only two types it follows that the only kind of bridged, sort-connected network is one where two agents of each of the two types break a link and form new links across. The threshold for optimality for this bridged, sort-connected can be computed from Eq. (A.24).

In order to prove properties (ii) and (iii) we want to show there are only two classes of networks which can be efficient: the sort-connected and the bridged, sort-connected. We begin by noting that utility under hyperbolic decay (from Equation 3) can be expressed as:

$$w_{ij}(\mu) = (1-\delta)\mathbf{1}_{=1}(p_{ij}(\mu)) + \delta \cdot \mathbf{1}_{\in[1,\infty)}(p_{ij}(\mu)). \tag{A.25}$$

Thus total utility from the network has the following form:

$$U(\mu) = \sum_{i\in N}\sum_{j\in N, j\neq i}\left[(1-\delta)\cdot\mathbf{1}_{=1}(p_{ij}(\mu)) + \delta\cdot\mathbf{1}_{\in[1,\infty)}(p_{ij}(\mu))\right]\cdot z(x_i,x_j). \tag{A.26}$$

The form for aggregate utility in Equation A.26 has the advantage that it is easier to perform optimization on. From inspection we see that if a network is connected then indirect term in the weights, $\delta \cdot \mathbf{1}_{\in[1,\infty)}(p_{ij}(\mu))$, is one for all edges, and as a consequence the aggregate utility attains its maximal value.

We first restrict ourselves to look at perfectly sorted networks. If it holds that each subnetwork $\mu_x \subseteq \mu$ consisting of all links within a given type is connected then the argument made above, that the aggregate utility from indirect links (i.e. stemming from $\delta \cdot \mathbf{1}_{\in[1,\infty)}(p_{ij}(\mu_x)) = 1$ for $x_i = x_j, i \neq j$ in Equation A.26), is maximized (conditional on perfect sorting). Finally, it must be that each subnetwork has no link surplus. This follows as there exists a subnetwork $\tilde{\mu}_x$ with no link surplus which is connected from Lemma 2. Thus any link surplus would imply inefficiency of μ_x as it would hold that the number of links between type x would be lower than the possible, i.e. $\sum_{ij\in\tilde{\mu}_x} \mathbf{1}_{=1}(p_{ij}(\tilde{\mu}_x)) > \sum_{ij\in\mu_x} \mathbf{1}_{=1}(p_{ij}(\mu_x))$, and thus provide lower welfare by Equation A.26. As any network with sorted connectivity obtains exactly the same utility we know that this set constitutes the set of networks with highest aggregate utility among networks with perfect sorting. We know from Proposition 3 that the set of networks with sorted connectivity is non-empty. We have thus shown that the set of networks with sorted connectivity are efficient among perfectly sorted networks.

We proceeed with analyzing efficient networks among those without perfect sorting. Assume that a network μ is not perfectly sorted. Suppose further that μ is connected. Then the total utility from indirect links is maximized as $\mathbf{1}_{\in[1,\infty)}(p_{ij}(\mu)) = 1$ for every $i \neq j$. The utility accruing from (direct) links stems from the term $\mathbf{1}_{=1}(p_{ij}(\mu))$ in Equation A.26. Due to supermodularity the utility from (direct) linking will be maximized if there is perfect sorting, however, this is not feasible as we require links across the two types. The minimal required links across types are two for every type. This follows as at least one link across types is required and thus the number of same type links must be at least one lower. Therefore, the highest attainable number of links within same type is $\frac{n_x\kappa}{2} - 1$ with two links across. Having $\frac{n_x\kappa}{2} - 1$ same type links and two cross-type links as well as type connectivity correspond exactly to the definition of bridged, sort-connected networks. Any other network which is not perfectly sorted can also at most have $\frac{n_x\kappa}{2} - 1$ same type links. This implies that the bridged, sort-connected has maximal benefits possible from direct (links) subject to being perfectly sorted. Due to being connected it also has the maximum number of indirect benefits. It remains to show that the set of bridged, sort-connected networks is non-empty; we can construct a bridged, sort-connected network from a sort-connected network as $\hat{\mu} = \hat{\mu} \cup \{ij, i'j'\} \backslash \{ii', jj'\}$ where $x_i = x_{i'}, x_j = x_{j'}, x_i \neq x_j$ and μ is sort-connected. We verify the that by construction $\hat{\mu}$ the has the feature of being connected (as the subnetworks for each type are connected if we choose each subnetwork using Lemma 2) and there are exactly $\frac{n_x\kappa}{2} - 1$ links of same type links for each type. Thus, we have determined that the bridged, sort-connected network must maximize aggregate utility among networks that are not-perfectly sorted.

We have established there are only two networks which can be efficient. We know from Eq. (A.24) that if $\delta < \delta^{opt}$ then the payoff from sort-connected network exceeds the payoff from bridged, sort-connected networks and vice versa. Therefore, when $\delta \leq \delta^{opt}$ then the sort-connected network is efficient, however, when $\delta \geq \delta^{opt}$ then the bridged, sort-connected network is efficient. ■

Proof Proposition 4.. Let $\mu \in \hat{M}$ and $\delta < \delta^{stab}$. By construction there exists a network $\tilde{\mu}$ which has higher aggregate utility. Let the two pairs of agents ii', jj' be agents such that $\tilde{\mu} = \mu \cup \{ij, i'j'\} \backslash \{ii', jj'\}$ and $x_i = x_{i'} = x$ and $x_j = x_{j'} = x$. Specify a link-contingent contract to i, j where $\hat{\mu} = \mu \cup \{ij\} \backslash \{ii', jj'\}$ such that:

$$\forall \iota\iota' \in \{ij, i'j'\}: \qquad \mathcal{C}_{\iota\iota'} + \mathcal{C}_{\iota'\iota} \geq \tfrac{1}{2}[Z(x,x) + Z(\tilde{x},\tilde{x}) - 2Z(x,\tilde{x})], \qquad \text{(A.27)}$$

$$\forall \iota\iota' \notin \{ij, ji, i'j', j'i'\}: \qquad \mathcal{C}_{\iota\iota'} = 0. \qquad \text{(A.28)}$$

By Theorem 1 we know that μ is pairwise stable. Pairwise stability implies that $\frac{1}{2}[Z(x,x) + Z(\tilde{x},\tilde{x}) - 2Z(x,\tilde{x})] > u_i(\mu) - u_i(\hat{\mu}) + u_j(\mu) - u_j(\hat{\mu})$ as deviation is not profitable. Using this fact together with Inequality A.27 it follows that:

$$\mathcal{C}_{ij} + \mathcal{C}_{ji} > u_i(\mu) - b_i(\hat{\mu}) + u_j(\mu) - u_j(\hat{\mu}).$$

The above inequality entails agents i, j are a blocking coalition that can gain by deviating to $\hat{\mu}$; this blocking move is also the only profitable move for i, j due to pairwise stability of μ and Equation A.28.

In network $\hat{\mu}$ agents i', j' have an incentive to form a link with one another as both have surplus link capacity (i.e. degree below the quota) and forming a link is profitable from Inequality A.27. Moreover, we show in the following that this move is the one that ensures the highest aggregate net benefits to i', j'.

We begin with showing that linking across types to other agents of type $x, \tilde{x}$ is not profitable. Suppose i' links across types to another agent $j'' \in \{\iota \neq j' : x_\iota = x_{j'}\}$. First, note the pairwise deviation from μ to form $i'j''$ is unprofitable (due to pairwise stability), thus it less profitable than forming $i'j'$ from μ (which is profitable by Inequality A.27). Second, the net-increase in value of the pairwise deviation to form $i'j'$ over $i'j''$ increases from μ to $\hat{\mu}$ - this is true as j' loses the link with i from μ while j'' has an unchanged number - thus j' will have a weakly lower opportunity cost of deleting links in $\hat{\mu}$. The same argument can be applied to j' for $i'' \in \{\iota \neq i' : x_\iota = x_{i'}\}$.

We turn to showing that linking to other agents of same type (staying sorted) is not more profitable as well. Suppose i' and j' link to same types as themselves respectively, i.e. $i'' \in \{\iota \neq i' : x_\iota = x_{i'}\}$ and $j'' \in \{\iota \neq j' : x_\iota = x_{j'}\}$. Suppose $ii'' \in \mu$ then no feasible pairwise moves to same type can exist in $\hat{\mu}$ as the move can only involve deleting links; same is true if $jj'' \in \mu$. Thus instead we use $ii'', jj'' \notin \mu$. It must be that any pairwise deviation forming either ii'' or jj'' from μ is unprofitable (as μ is pairwise stable); this implies that for any $\iota \in \nu_{i''}(\hat{\mu})$ and $\iota' \in \nu_{j''}(\hat{\mu})$ it holds that:

$$u_{i'}(\hat{\mu} \cup \{i'i''\} \backslash \{i''\iota\}) - u_{i'}(\hat{\mu}) + u_{i''}(\hat{\mu} \cup \{i'i''\} \backslash \{i''\iota\}) - u_{i''}(\hat{\mu}) - \tau_{i''\iota} \leq z(x, x) \tag{A.29}$$

$$u_{j'}(\hat{\mu} \cup \{j'j''\} \backslash \{j''\iota'\}) - u_{j'}(\hat{\mu}) + u_{j''}(\hat{\mu} \cup \{j'j''\} \backslash \{j''\iota'\}) - u_{j''}(\hat{\mu}) - \tau_{j''\iota'} \leq z(\tilde{x}, \tilde{x}) \tag{A.30}$$

As $u_{i'}(\tilde{\mu}) - u_{i'}(\hat{\mu}) + u_{j'}(\tilde{\mu}) - u_{j'}(\hat{\mu}) = z(x, \tilde{x}) + z(\tilde{x}, x)$ it follows that

$$u_{i'}(\tilde{\mu}) - u_{i'}(\hat{\mu}) + u_{j'}(\tilde{\mu}) - u_{j'}(\hat{\mu}) + \mathcal{C}_{i'j'} + \mathcal{C}_{j'i'} > z(x, x) + z(\tilde{x}, \tilde{x}).$$

The above inequality implies together with Inequalities A.29 and A.30 that the total gains for i' and j' exceeds the total value that could be generated from alternative deviations. Thus there are two pairwise moves from μ to $\hat{\mu}$ and from $\hat{\mu}$ to $\tilde{\mu}$ which both provide strictly higher utility to the deviating agents.

Pairwise stability follows from three arguments. First, all deviations among agents where only links in $\tilde{\mu} \cap \mu$ are deleted will provide at most the same value in $\tilde{\mu}$ that the deviations did in μ - this follows as these agents all have the same links and in $\tilde{\mu}$ all agents are connected in $\tilde{\mu}$ and thus only direct links matter. This upper limit too gains from deviations implies none of these moves can be profitable as they were unprofitable form μ. Second, deviations that involve deletion of links in $\tilde{\mu} \backslash \mu$ are shown above to provide strictly higher value than any other deviations - thus deviating from $\tilde{\mu}$ must also provide strictly lower value. ■

Proof of Proposition 5.. Under asymptotic independence it follows that average per agent utility for type x under asymptotic perfect sorting converges to (using a geometric series):

$$\frac{(\kappa - 1)\,\delta}{1 - (\kappa - 1)\,\delta} z(x, x)$$

Let $\omega_{x\tilde{x}} = \kappa \cdot \mathbb{E}[\delta^{p_{ij}} | x_i = x, x_j = \tilde{x}]$. Suppose that for two types, $x, \tilde{x}$ there is not perfect sorting, and in particular there is some mixing between them, i.e. $\omega_{x\tilde{x}} > 0$; the average per agent utility is:

$$\left[\frac{(\kappa - 1)\,\delta}{1 - (\kappa - 1)\,\delta} - \omega_x \right] \cdot z(x, x) + \omega_x \cdot z(x, \tilde{x}).$$

Each agent will almost surely have κ links as it is assumed that each link adds positive value and there are asymptotic infinite agents (only a finite number can then not fulfill the degree quota).

As we have a finite set of types we can assume then for large populations there is a subset of types, $\hat{X} \subseteq X$, where for every type $x \in \hat{X}$ it holds that there is an asymptotic strictly positive share of the total number of agents of that type, i.e., $\lim_{n\to\infty}(|\{i \in N_n\}_{x_i=x}|/n) > 0$. If there is only one such type, i.e. $|\hat{X}| = 1$, then asymptotic perfect sorting follows by assumption as the asymptotic number of links is κ.

For any two types $x, \tilde{x} \in \hat{X}$ which are mixing their average utility is:

$$\frac{(\kappa-1)\,\delta}{1-(\kappa-1)\,\delta}\left[\frac{n_x \cdot z(x,x) + n_{\tilde{x}} \cdot z(\tilde{x},\tilde{x})}{n_x + n_{\tilde{x}}}\right] - \frac{1}{2}\cdot\left[\frac{n_x \cdot \omega_{x\tilde{x}}}{n_x + n_{\tilde{x}}}\right]\cdot[Z(x,x) + Z(\tilde{x},\tilde{x}) - 2Z(x,\tilde{x})].$$

As there is supermodularity it follows that $Z(x,x) + Z(\tilde{x},\tilde{x}) - 2Z(x,\tilde{x}) > 0$ and thus mixing must decrease utility. The same argument can be applied by mixing between multiple types. ■

Appendix B. Supplementary appendix: finite poulation and constant decay in spillovers

This appendix extends the analysis of sorted networks with finite number of agents to a setting where decay is constant. We prove properties of stability and efficiency for a sub-class of sort-connected networks under constant decay. Specifically we show that certain network with sorted connectivity are pairwise stable for low to moderate spillovers. We also show that these network are suboptimal for moderate to high spillovers.

The appendix is split into two sub-appendices, Appendix B.1 which contains the main results and Appendix B.2 which only contains auxiliary results.

Appendix B.1. Suboptimal sorting in local trees

We begin by describing the sub-class of sort-connected networks. Informally put, the sub-class has the added requirement that networks are not only connected within each type, but also resembles a certain tree structure. We define a *tree* as a network where every pair of agents are connected by a unique path. The structure of each subnetwork is such that from the perspective of every agent (i.e. the ego-network) each subnetwork appears as a tree locally. That is, the network becomes a tree if we remove all links for the agents furthest away from the considered agent which are not on their shortest path(s) to the considered agent. Therefore the networks have a local tree structure but not a global one.

The formal definition of local trees is described below. The definition employs the network *diameter* which is the maximum distance between any two agents, i.e. $m(\mu) = \sup_{i,j\in N} p_{ij}(\mu)$.

Definition 11. *A network μ is a **local tree** when each agent i has κ links where:*

- *for each other agent $j \neq i$ at distance $p_{ij}(\mu) \leq m_{n,\kappa} - 2$ there are $\kappa - 1$ links between agent j and j' such that j' is one step further away, i.e. $p_{ij}(\mu) = p_{ij'}(\mu) - 1$;*
- *the network diameter $m(\mu) = m_{n,\kappa}$,*

$$m_{n,\kappa} \quad = \quad \arg\min_{m}\{m : \Sigma_{l=1}^{m}(\kappa(\kappa-1)^{l-1}) + 1 \geq n\}. \qquad \text{(B.1)}$$

The structure of local trees entails that each agent has $\kappa \cdot (\kappa - 1)^{p-1}$ agents at distance $p < m$, where $m = m_{n,\kappa}$. At distance $p = m$ there are $n - \sum_{l=1}^{m-1} \kappa \cdot (\kappa - 1)^{l-1}$ (all remaining agents). This structure implies that every agent's utility is maximized subject to the constraint of all agents having at most κ links;[20] a side effect is that utility before transfers is symmetric.

A necessary condition for local trees to exist is that there is no link surplus, i.e. degree quota is binding ($\forall i \in N : k_i = \kappa$). Note this binding condition is only possible when $n \cdot \kappa$ is even.

When a local tree network fulfills $n = \sum_{l=1}^{m} \kappa \cdot (\kappa - 1)^{l-1}$ then it is an *exact local tree*. See the next sub-appendix for an elaborate treatment of structure of exactly local trees. Two subclasses of exact local trees which are worth mentioning. The first is a network known as a cycle or a ring. The cycle is characterized by having a minimal possible degree quota ($\kappa = 2$) among local trees and a maximal diameter ($m = \left\lceil \frac{n-1}{2} \right\rceil$). The second is a *clique* where all agents are linked, i.e. the complete network. Cliques have maximal degree quotas ($\kappa = n - 1$) and minimal diameters ($m = 1$). Both subclasses has a network which exists for any n. Note that in Example 1 each of the two components is both a cycle and a clique. Note that there exist non-trivial networks beyond the cycle and the clique.[21]

---

[20]The maximization of utility follows from the observation that each agent has at most κ links, so at distance p there can be at most $\kappa \cdot (\kappa - 1)^{p-1}$ agents.

[21]An example is $\{i_1i_2, i_1i_3, i_1i_4, i_2i_5, i_2i_6, i_3i_7, i_3i_8, i_4i_9, i_4i_{10}, i_5i_7, i_5i_9, i_6i_8, i_6i_{10}, i_7i_9, i_8i_{10}\}$ when $n = 10, \kappa = 3$ and $N = \{i_1, i_2, .., i_{10}\}$.

In order to derive our results it is necessary to restrict ourselves to a subset of local trees. The subset are those local trees where the deletion of links leads to equal losses to both of agents whose link is deleted; thus we refer to these local trees as having symmetric losses:

Definition 12. *A local tree μ has* ***symmetric losses*** *when at every distance $p = 1,..,m$ it holds that $|\{i \in N : p_{\iota i}(\mu \backslash \{\iota\iota'\}) = p\}| = |\{i \in N : p_{\iota' i}(\mu \backslash \{\iota\iota'\}) = p\}|$.*

Whether or not symmetric losses is a generic property for all local trees is an open question. However, in simulations that we perform it holds all network configurations which are local trees up to size $n = 10$ have symmetric losses (see result below and proof for exhibition of examples). In addition, for size up to $n = 16$ it has been shown to hold for any networks examined in the simulation.

A generalization of stable but suboptimal sorting under constant decay is expressed below. While allowing for constant decay rather than hyperbolic it the set of networks are further restricted.

Theorem 3. *Suppose there is supermodularity, a degree quota κ and each type has equal number of agents. It follows that any network which is sort-connected and consist of symmetric local trees is also: (i) pairwise stable if $\delta < \delta^{stab}$; and, (ii) inefficient if $\delta > \delta^{opt}$. The thresholds satisfy $\delta^{opt}, \delta^{stab} \in (0, 1)$ where $\delta^{opt} < \delta^{stab}$*

Proof. We show properties (i) and (ii) together. Let μ be a network which is perfectly sorted into $|X|$ components, one for each type. Each component is a local tree with $n/|X|$ agents. Let there be no transfers between any agents.

As each subnetwork for a given type is a local tree it is stable against deviations by agents of the same type - this follows as local trees provides maximal possible benefits among feasible structures of the subnetwork for all agents in the subnetwork. Thus only two agents of different types may have a profitable deviation which is feasible.

Let ι, j be agents of respectively types x and $\tilde{x}$. These two agents can deviate by each deleting a link to ι' and j' respectively while jointly forming a link. The new network resulting from deletion is denoted $\hat{\mu} = \mu \backslash \{\iota\iota', jj'\}$. The move resulting from deletion and forming a link is denoted $\breve{\mu} = \hat{\mu} \cup \{\iota j\}$. An alternative network is $\tilde{\mu}$, the type-bridged network of μ, where the links $\iota\iota', jj'$ are removed while the links $\iota j, \iota', j'$ have been formed; thus $\tilde{\mu} = \hat{\mu} \cup \{\iota j, \iota' j'\}$.

Define the gross loss of benefits for i as $u_i(\hat{\mu}) - u_i(\mu)$ while the gross gains are $u_i(\tilde{\mu}) - u_i(\hat{\mu})$. There must exist a threshold of externalities $\delta^{stab} \in (0, 1)$ where μ is no longer pairwise stable as cost of deviation monotonically decreases and approaches zero as $\delta \to 1$ while gains are monotonically increasing. The monotonicity of losses is a consequence of the fact that gross loss consists of shortest paths from μ, where $\iota\iota'$ is included in the shortest path, which have longer length in $\hat{\mu}$ and thus are discounted more. Therefore the gross loss is mitigated by a higher δ as the longer shortest paths are punished less. The monotonicity of gains follows as the gains consist of new shortest paths to agents of type $\tilde{x}$ through ιj and $j'\iota'$ the value of these increases for higher δ.

Exploiting the that Fact 1 and 2 from Appendix Appendix B.2 hold for local trees it follows that for any other agent i of type x (i.e. i is in $N \backslash \{\iota, \iota'\}$ and $x_i = x$):

$$u_i(\tilde{\mu}) - u_i(\mu) > \delta^{\min(p_{i\iota}(\tilde{\mu}), p_{i\iota'}(\tilde{\mu}))}[u_\iota(\breve{\mu}) - u_\iota(\mu)].$$

Aggregating for all agents this implies:

$$U(\tilde{\mu}) - U(\mu) \;\; > \;\; [u_\iota(\breve{\mu}) - u_\iota(\mu)] \cdot \sum_{x_i = x} \delta^{\min(p_{i\iota}(\tilde{\mu}), p_{i\iota'}(\tilde{\mu}))} + [u_j(\breve{\mu}) - u_j(\mu)] \cdot \sum_{x'_i = \tilde{x}} \delta^{\min(p_{ij}(\mu), p_{ij'}(\mu))}.$$

where $m = m_{n,\kappa}$. The inequality above implies the following: if $U(\tilde{\mu}) - U(\mu) = 0$ then $u_\iota(\breve{\mu}) - u_\iota(\mu) + u_j(\breve{\mu}) - u_j(\mu) < 0$; $U(\tilde{\mu}) - U(\mu) > 0$ when $u_\iota(\breve{\mu}) - u_\iota(\mu) + u_j(\breve{\mu}) - u_j(\mu) = 0$. It can also be argued that there must exist a threshold, δ^{opt}, such that when $\delta = \delta^{opt}$ then $U(\tilde{\mu}) - U(\mu) = 0$ and that $\delta^{opt} < \delta^{stab}$. This follows as $U(\tilde{\mu}) - U(\mu) < 0$ for $\delta = 0$ and $U(\tilde{\mu}) - U(\mu) > 0$ when $u_\iota(\tilde{\mu}) - u_\iota(\mu) + u_j(\tilde{\mu}) - u_j(\mu) = 0$ as well as continuity of $U(\tilde{\mu}) - U(\mu)$ in δ.

This entails that for $\delta > \delta^{opt}$ then $\tilde{\mu}$ provide higher aggregate payoff. Moreover we showed previously that for $\delta < \delta^{stab}$ then μ is pairwise (Nash) stable. Thus we have proven properties (i) and (ii).

□

For constant decay the thresholds governing when sorting is respectively suboptimal and stable, i.e. $\delta^{opt}, \delta^{stab}$, can be determined explicitly by solving polynomial equations for every deviation. Moreover, for exact local trees there is a unique solution. In Figure B.3 the two thresholds from Theorem 3, $\delta^{opt}(\hat{Z}), \delta^{stab}(\hat{Z})$.

The plots in Figure B.3 are made for variations of exact local trees. The upper plots corresponds to cliques with various sizes. The lower plot have fixed degree quota (κ=100) and the threshold is simulated using pattern in utility that is demonstrated in Appendix Appendix B.2. The plots show the scope for inefficiency, i.e. the gap between $\delta^{opt}(\hat{Z}), \delta^{stab}(\hat{Z})$, increases with the number of agents involved. This makes sense intuitively as the two agents forming the link will fail to account for an increasing number of indirect connections between the two groups. As the number of indirect connections increases at with the squared with total number of agents then larger populations will lead to larger gaps of inefficiency.

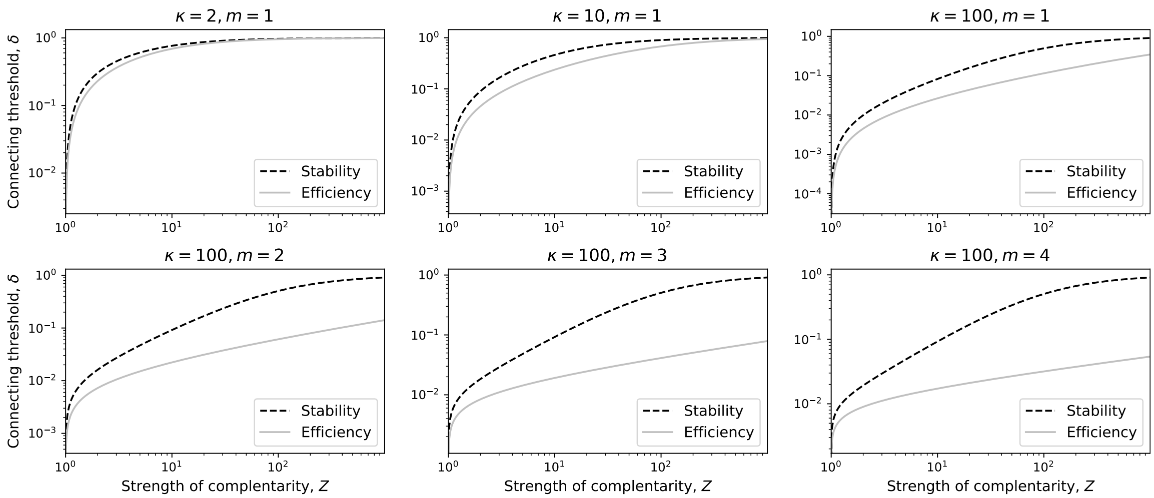


Figure B.3: *Visualization of thresholds for connecting from Theorem 3.*

The upper diagrams correspond to cliques and the lower ones to exact local trees (where thresholds stem from Equations B.4, B.5, B.10, B.11).

Appendix B.2. Local trees

This sub-appendix provides auxiliary results for deriving the generalization of suboptimal sorting. We begin our focus on exact local trees and subsequently more generally in local tree networks, see Definition 11 in the previous sub-appendix.

We will examine a generic network μ which is perfectly sorted and assume that the subset of links for each type is a component that can be classified as either a local tree or an exact local tree. Let networks μ_x and $\mu_{\tilde{x}}$ be the components associated with respectively types $x, \tilde{x} \in X$. We will focus on three particular moves:

- *Pairwise deletion of a link*: Suppose two links $\iota\iota', jj' \in \mu$ are deleted and agents ι and j have respectively type x and $\tilde{x}$; thus the two links are not from the same component. Let the new network that results from removal of the links be denoted $\hat{\mu} = \mu \backslash \{\iota\iota', jj'\}$.

- *Pairwise formation of a link across types*: This move presumes that both agents are also deleting a link. We denote this as a move where agents ι and j form a link: $\breve{\mu} = \hat{\mu} \cup \{\iota j\}$.

- *Double pairwise formation of a link across types*: When two links are formed across types in μ this corresponds to a non-pairwise deviation as it requires four coalition members. We denote this as a move where both agents ι and j as well as ι' and j' form a link: $\tilde{\mu} = \hat{\mu} \cup \{\iota j, \iota' j'\}$.

Finally let i denote a generic agent of type x. Let the shortest path in μ from i to either ι or ι' be denoted $\hat{p}_i$ where $\hat{p}_i = \min(p_{i\iota}(\hat{\mu}), p_{i\iota'}(\hat{\mu}))$. When $\hat{p}_i = 0$ then either $i = \iota$ or $i = \iota'$.

Basic properties. We exploit that μ is a local tree (see Definition 11). Throughout the remainder of the paper let $m = m_{n,\kappa}$ (see Equation B.1). We express each agent's number of paths of length p as a function of the number of agents and the degree quota:

$$\#_i^p(\mu) = \kappa(\kappa-1)^{p-1} - \mathbf{1}_{=m}(p) \cdot \Delta\#(n,\kappa), \qquad \Delta\#(n,\kappa) = \sum_{l=1}^{m} (\kappa \cdot (\kappa-1)^{l-1}) - n, \quad \text{(B.2)}$$

where $\mathbf{1}_{=m}(p)$ is the Dirac measure of whether $p = m$. Using the local tree structure we can express utility without transfers of each agent:

$$u_i(\mu) = \sum_{l=1}^{m} \#_i^l(\mu) \cdot \delta^l \cdot z(x,x).$$

Exact local trees

Recall exact local trees are local trees where $\Delta\#(n,\kappa) = 0$. We will argue that this entails that exact local trees have the essential property that for every pair of agents there is a unique shortest path of at most length m and the number of paths for every agent is prescribed by Equation B.2. This can be deducted as follows.

Note first that the fact that the number of walks with at most length m starting in a given agent i cannot exceed $\sum_{p=1}^{m} \#_i^p(\cdot)$. Recall also that local trees has the property that all agents are reached within distance m. Moreover exact local trees has the property that for any agent i it holds that $n - 1 = \sum_{p=1}^{m} \#_i^p(\mu)$; thus every shortest path with distances less than or equal to m must be a unique path between the two particular agents.

The uniqueness and countability of paths can be used to infer the losses when links are either removed or added to an exact local tree.

Exact local trees - loss from deletion. In order to examine the impact of deletion of a link it is sufficient to analyze what happens to one component of types. This is sufficient as other components as the conclusions are valid for all.

The deletion of link $\iota\iota'$ implies that any pair of agents i, i' whose (unique) shortest path in μ includes the link $\iota\iota'$ will have a new shortest routing path. For exact local trees we can exactly determine the length of the new path. Let i be the agent whose distance to ι is least and let i' be the agent whose distance to ι' is least, i.e. $p_{i\iota}(\mu) < p_{i\iota'}(\mu)$ and $p_{i'\iota'}(\mu) < p_{i'\iota}(\mu)$.

First when link $\iota\iota'$ is deleted we can show there is no shortest path between i and i' in $\hat{\mu}$ with length below $2m - \hat{p}_i - \hat{p}_{i'}$; that is there is no ii' whose shortest path in μ includes $\iota\iota'$ such that $p_{ii'}(\hat{\mu}) < 2m - \hat{p}_i - \hat{p}_{i'}$. Suppose this was not true. Then there would exist an agent j who (1) is on the new shortest path between i and i' in $\hat{\mu}$ and (2) whose shortest path to agents ι and ι' does not include the link $\iota\iota'$ and (3) such that

$$\begin{aligned} p_{ji}(\hat{\mu}) + p_{ji'}(\hat{\mu}) &< 2m - \hat{p}_i - \hat{p}_{i'}, \\ p_{ji}(\hat{\mu}) + p_{ji'}(\hat{\mu}) &< 2m - \min(p_{i\iota}(\mu),\, p_{i\iota'}(\mu)) - \min(p_{i'\iota}(\mu),\, p_{i'\iota'}(\mu)). \end{aligned}$$

As by construction $p_{i\iota}(\mu) < p_{i\iota'}(\mu)$ and $p_{i'\iota'}(\mu) < p_{i'\iota}(\mu)$ then the expression above is equivalent to: $p_{ji}(\hat{\mu}) + p_{ji'}(\hat{\mu}) < 2m - p_{i\iota}(\mu) - p_{i'\iota'}(\mu)$. As the shortest path between i and ι as well as between i' and ι' are unchanged from μ to $\hat{\mu}$ it follows that we can further rewrite into:

$$p_{ji}(\hat{\mu}) + p_{ji'}(\hat{\mu}) < 2m - p_{i\iota}(\hat{\mu}) - p_{i'\iota'}(\hat{\mu})$$

However, the above statement implies that in network μ that either ι or ι' has two paths with lengths of at most m but this violates the definition of exact local trees.

We can now show that when link $\iota\iota'$ is deleted the new shortest path between i and i' in $\hat{\mu}$ has a length of exactly $2m - \hat{p}_i - \hat{p}_{i'}$. This is shown by demonstrating there is an agent j such that $p_{ji}(\hat{\mu}) = m - \hat{p}_i$ and $p_{ji'}(\hat{\mu}) = m - \hat{p}_{i'}$. This can be shown follows. Suppose that $p_{ji}(\hat{\mu}) = m - \hat{p}_i$. We will demonstrate that $p_{ji'}(\hat{\mu}) = m - \hat{p}_{i'}$. As $p_{ji}(\hat{\mu}) = m - \hat{p}_i$ it follows that $p_{j\iota}(\hat{\mu}) = m$. From the definition of exact local trees there must exist a path of length less than m between j and ι' in network μ. As argued in the paragraph above neither of these paths can be strictly shorter than m and consequently they must both be exactly m.

The number of shortest paths of length p which become altered for agent i is $(\kappa - 1)^{p - \hat{p}_i - 1}$ for $p = \hat{p}_i, .., m-2, m-1$. This can be demonstrated as follows. If agent $p_{i\iota}(\mu) = m$ and $p_{i\iota'}(\mu) = m$ then no shortest paths are altered; this is clear as agent i as none of the unique shortest paths includes $\iota\iota'$ as they have at most length m. If instead $p_{i\iota}(\mu) = m - 1$ then the unique shortest path from i to ι' includes $\iota\iota'$ is the last link; this implies a new shortest path if $\iota\iota'$ is deleted. Thus if $p_{i\iota}(\mu) = m - 1$ then one shortest path of length m is lost. When $p_{i\iota}(\mu) = m - 2$ then one path of length $m-1$ is lost by the same argument; moreover $\kappa-1$ paths that has $\iota\iota'$ as the second last link. By induction this can be done at higher order and thus for shorter distances. Using the number of rerouted paths shown above we can establish the total number of shortest paths in network $\hat{\mu}$ for agent i that has a length of p:

$$\#_i^p(\hat{\mu}) = \begin{cases} \kappa(\kappa-1)^{p-1} - \mathbf{1}_{>\hat{p}_i}(p) \cdot (\kappa-1)^{p-\hat{p}_i-1}, & p \le m \\ (\kappa-1)^{2m-\hat{p}_i-p}, & p \in (m, 2m-\hat{p}_i]. \end{cases} \tag{B.3}$$

By combining the count of shortest paths rerouted with their new length we can generalize the loss for any agent from the deletion of link $\iota\iota'$ when all agents are homogeneous of type x:

$$u_i(\mu) - u_i(\hat{\mu}) = \sum_{l=1}^{m-\hat{p}_i} \left[(\kappa-1)^{l-1} \cdot \left(\delta^{l-1+\hat{p}_i} - \delta^{2m-(l-1)-\hat{p}_i}\right)\right] \cdot z(x,x). \tag{B.4}$$

We can aggregate the losses across homogeneous agents of type x and we arrive at the following expression:

$$U(\mu) - U(\hat{\mu}) = \sum_{l=1}^{m} \left[2l \cdot (\kappa-1)^{l-1} \cdot \left(\delta^{l-1} - \delta^{2m-(l-1)}\right)\right] \cdot z(x,x). \tag{B.5}$$

Exact local trees - gains from linking across types. We move on to establishing the gains of establishing a link in a perfectly sorted network where each component is an exact local tree.

The gains to agents ι and j of forming a link ιj are direct benefits and the new indirect connections that are accessed through the link ιj. For agent ι the benefits from forming a link with j can be computed with Equation B.3 where the input length is added one (as ιj is added to the shortest path). Recall $\breve{\mu} = \mu \cup \{\iota j\} \backslash \{\iota\iota', jj'\}$.

$$u_\iota(\breve{\mu}) - u_\iota(\mu) = \left[\sum_{l=0}^{m}(\kappa-1)^l \cdot \delta^l + \sum_{l=0}^{m-1}(\kappa-1)^l \cdot \delta^{2m-l}\right] \cdot z(x,\tilde{x}). \tag{B.6}$$

The above expression is relevant for evaluating the pairwise gains as it captures individual benefits for a pairwise formation of a link by ι and j. However, we are also interested in the sub-connected network as it allows to assess the efficiency. Suppose instead now that ι' and j' also form a link; thus $\iota j, \iota' j'$ are formed while $\iota\iota', jj'$ are deleted. Let $\tilde{\mu} = \mu \cup \{\iota j, \iota' j'\} \backslash \{\iota\iota', jj'\}$.

Let i be an agent of type x and let $\hat{p}_i$ still denote the least distance to either ι or ι'. We can calculate the benefits for i when $\iota j, \iota' j'$ are formed. The benefits are the indirect connections to agents of type $\tilde{x}$ with whom agent i has no connections in μ. The aim is to count the number of paths of a given length.

For a given agent i' of the other type $\tilde{x}$ it must hold that the shortest path in $\tilde{\mu}$ between i, i' either contains the link ιj or the link $\iota' j'$, and thus the distance can be computed as follows:

$$p_{ii'}(\tilde{\mu}) \;=\; \min[p_{ij}(\tilde{\mu}) + p_{i'j}(\tilde{\mu}),\; p_{ij'}(\tilde{\mu}) + p_{i'j'}(\tilde{\mu})] \tag{B.7}$$

We further restrict the above expression. We can use that i and i' of type $\tilde{x}$ can be at most $2m+1$ away from each other. This follows from the fact that $p_{i\iota}(\tilde{\mu}) + p_{i\iota'}(\tilde{\mu}) = 2m$ and $p_{i'j}(\tilde{\mu}) + p_{i'j'}(\tilde{\mu}) = 2m$. As $p_{i\iota}(\tilde{\mu}) + p_{i\iota'}(\tilde{\mu}) = 2m$ and $\iota j, \iota' j' \in \tilde{\mu}$ then it must be that $p_{ij} + p_{ij'} = 2m+2$. These facts together entail we can rewrite Equation B.7:

$$\begin{aligned} p_{ii'}(\tilde{\mu}) &= \min[p_{ij}(\tilde{\mu}) + p_{i'j}(\tilde{\mu}),\; p_{ij'}(\tilde{\mu}) + p_{i'j'}(\tilde{\mu})] \\ &= \min[p_{ij}(\tilde{\mu}) + p_{i'j}(\tilde{\mu}),\; 4m+2 - p_{ij}(\tilde{\mu}) - p_{i'j}(\tilde{\mu})]. \end{aligned} \tag{B.8}$$

From the above expression it follows that $p_{ii'} \leq 2m+1$ as the expression is maximized for $p_{ij} + p_{i'j} = 2m+1$.

The number of shortest paths from i through ιj to agents of the other type $\tilde{x}$ can be found using Equation B.3 for agent ι adding extra distance $1 + \hat{p}_i$:[22]

- for distance $p \in \{1 + \hat{p}_i, ..., m+1+\hat{p}_i\}$ there are $(\kappa-1)^{p-1-\hat{p}_i}$ agents;
- for distance $p \in \{m+2+\hat{p}_i, ..., 2m+1\}$ there are $(\kappa-1)^{2m+1-(p-1-\hat{p}_i)}$.

The shortest paths from i not routed through ι but instead through ι' are those where $p+1+\hat{p}_i > 2m+1$; from Equation B.8 we know the new shortest path length is $4m+2-p-1-\hat{p}_i$. The number of shortest paths through ι' in network $\tilde{\mu}$ will be $(\kappa-1)^{2m+1-(p-1-\hat{p}_i)}$ and the new length $4m+2-p-1-\hat{p}_i$. These facts together imply:

$$\#_i^p(\tilde{\mu}) - \#_i^p(\hat{\mu}) = \begin{cases} (\kappa-1)^{p-1-\hat{p}_i}, & p \in \{\hat{p}_i+1, .., m+1+\hat{p}_i\}, \\ (\kappa-1)^{2m+1-p-\hat{p}_i}, & p \in \{m+\hat{p}_i+2, .., 2m+1\}, \\ (\kappa-1)^{p+\hat{p}_i-2m-1}, & p \in \{2m+1-\hat{p}_i, .., 2m\}. \end{cases} \tag{B.9}$$

[22] Shortest paths from i must contain both ιj and every link in the shortest path from i to j.

From the number of paths above we can derive the change in utility from when $\iota j, \iota' j'$ are added to the network for a given agent i of type x.

$$u_i(\tilde{\mu}) - u_i(\hat{\mu}) = \begin{bmatrix} & \sum_{l=0}^{m} (\kappa-1)^l \cdot \delta^{l+\hat{p}_i} \\ + & \sum_{l=\hat{p}_i}^{m-1} (\kappa-1)^l \cdot \delta^{2m-l+\hat{p}_i} \\ + & \sum_{l=0}^{\hat{p}_i-1} (\kappa-1)^l \cdot \delta^{2m+l-\hat{p}_i} \end{bmatrix} \cdot z(x, \tilde{x}). \tag{B.10}$$

By aggregating over all agents of type the gain in benefits by forming $\iota j, \iota' j'$ is as follows:

$$U(\tilde{\mu}) - U(\hat{\mu}) = \sum_{p=0}^{m} \left(\begin{bmatrix} \mathbf{1}_{<m}(p) \cdot 2 \cdot (\kappa-1)^p + \\ \mathbf{1}_{=m}(p) \cdot (n - 2 \cdot \sum_{l=1}^{m-1} (\kappa-1)^l) \end{bmatrix} \cdot \begin{bmatrix} & \sum_{l=0}^{m} (\kappa-1)^l \cdot \delta^{l+p} \\ + & \sum_{l=p}^{m-1} (\kappa-1)^l \cdot \delta^{2m-l+p} \\ + & \sum_{l=0}^{p-1} (\kappa-1)^l \cdot \delta^{2m+l-p} \end{bmatrix} \right) \cdot Z(x, \tilde{x}). \tag{B.11}$$

Local trees

We can use the analysis above on exact local trees to bound the gains and losses for (non-exact) local trees. Recall that exact local trees has the property that $\Delta\#(n, \kappa) = 0$ and for non-exact local trees $\Delta\#(n, \kappa) > 0$. Thus the difference between exact and non-exact local trees is that for a given agent the number of connected other agents at exactly distance m is lower for non-exact local trees.

Using the analysis of exact local trees we can compute the bounds on loss of utility for a given agent in the local when a link is deleted - this is done by reusing Equation B.3 as follows.

We can discount the number of agents initially at distance m by $\Delta\#(n, \kappa)$. Moreover, the new distance between agents i and i' after deletion of the link $\iota\iota'$ is at least $\min(p_{ii'}, 2m-2-\hat{p}_i-\hat{p}_{i'})$ at most $2m-\hat{p}_i-\hat{p}_{i'}$.[23] From these two facts we can derive the bound in loss of utility when $\iota\iota'$ is deleted. The upper bound of loss (in terms of magnitude) is found when new shortest paths have most distance, i.e. $2m - \hat{p}_i - \hat{p}_{i'}$; the lower bound is found when new distance is least, i.e. $\min(p_{ii'}, 2m-2-\hat{p}_i-\hat{p}_{i'})$:

[23] The upper bound follows from the fact that for any two agents i and i' in the local tree there is still always an agent j at distances $p_{ij} = m - \hat{p}_i$ and $p_{i'j} = m - \hat{p}_{i'}$. The lower bound can be established by repeating an argument used for exact local trees. If the new distance between two agents i and i' after deletion of $\iota\iota'$ had been less than $\min(p_{ii'}(\mu), 2m-2-\hat{p}_i-\hat{p}_{i'})$ then the following would be true. There would be multiple shortest paths of length less than or equal to $m-1$ between either (ι and j) or (ι' and j). This would violate the property of local trees that all shortest paths of length $\leq m-1$ are unique.

$$u_i(\mu) - u_i(\hat{\mu}) \;\leq\; \sum_{l=1}^{m-\hat{p}_i} \Big[\max(0, (\kappa-1)^{l-1} - \mathbf{1}_{=m}(l)\cdot \Delta\#(n,\kappa)) \left(\delta^{l-1+\hat{p}_i} - \delta^{2m-(l-1)-\hat{p}_i}\right)\Big] \cdot z(x,x), \tag{B.12}$$

$$u_i(\mu) - u_i(\hat{\mu}) \;\geq\; \sum_{l=1}^{\tilde{m}} \Big[(\kappa-1)^{l-1} \cdot \left(\delta^{l-1+\hat{p}_i} - \delta^{2m-(l+1)-\hat{p}_i}\right)\Big] \cdot z(x,x), \quad \tilde{m} = \min(m-1, m-\hat{p}_i). \tag{B.13}$$

Fact 1. *If μ is perfectly sorted and consists of $|X|$ components that each constitute a local tree with $n/|X|$ agents, then for any agent i of type x where $\hat{p}_i > 0$:*

$$u_i(\hat{\mu}) - u_i(\mu) > \delta^{\hat{p}_i} \cdot [u_\iota(\hat{\mu}) - u_\iota(\mu)], \quad \hat{p}_i = \min(p_{i\iota}(\hat{\mu}), p_{i\iota'}(\hat{\mu})). \tag{B.14}$$

Proof. Inequality B.14 can be rewritten into: $\delta^{\hat{p}_i} \cdot [u_\iota(\mu) - u_\iota(\hat{\mu})] - [u_i(\mu) - u_i(\hat{\mu})] > 0$. This inequality is equivalent to the expression below (derived by substituting in Inequality B.13 for agent ι and Inequality B.12 for agent i):

$$\delta^{\hat{p}_i} \cdot \sum_{l=1}^{m-1} \Big[(\kappa-1)^{l-1} \cdot \left(\delta^{l-1} - \delta^{2m-(l+1)}\right)\Big] - \sum_{l=1}^{m-\hat{p}_i} \Big[(\kappa-1)^{l-1} \left(\delta^{l-1+\hat{p}_i} - \delta^{2m-(l-1)-\hat{p}_i}\right)\Big] \;>\; 0,$$

$$\sum_{l=1}^{m-\hat{p}_i} \Big[(\kappa-1)^{l-1} \cdot \left(\delta^{2m-(l+1)-\hat{p}_i} - \delta^{2m-(l+1)+\hat{p}_i}\right)\Big] + \sum_{l=m-\hat{p}_i+1}^{m-1} \Big[(\kappa-1)^{l-1} \left(\delta^{l-1+\hat{p}_i} - \delta^{2m-(l-1)-\hat{p}_i}\right)\Big] \;>\; 0.$$

As it holds that $2m-(l+1)-\hat{p}_i < 2m-(l+1)+\hat{p}_i$ and it holds that $l-1+\hat{p}_i < 2m-(l-1)-\hat{p}_i$ (equivalent to $l < m+1-\hat{p}_i$) the above inequality is satisfied. □

We can also derive bounds on the gains from connecting across types for local trees. We will not do this explicitly but instead use Definition 12 on symmetric losses in local trees. This allows to express our next result:

Fact 2. *For the perfectly sorted network μ which consists of $|X|$ network components which each constitute a local tree of $n/|X|$ agents that has symmetric losses then it holds that for agents i, ι of type x and $\hat{p}_i > 0$*

$$u_i(\tilde{\mu}) - u_i(\hat{\mu}) \geq \delta^{\hat{p}_i} \cdot [u_\iota(\breve{\mu}) - u_\iota(\hat{\mu})], \quad \hat{p}_i = \min(p_{i\iota}(\hat{\mu}), p_{i\iota'}(\hat{\mu})). \tag{B.15}$$

Proof. It holds that $u_\iota(\tilde{\mu}) - u_\iota(\hat{\mu}) \geq u_\iota(\breve{\mu}) - u_\iota(\hat{\mu})$ as $\tilde{\mu} \subseteq \breve{\mu}$ (thus all shortest paths in $\tilde{\mu}$ cannot have a length that exceeds that in $\breve{\mu}$). Therefore it suffices to show:

$$u_i(\tilde{\mu}) - u_i(\hat{\mu}) \geq \delta^{\hat{p}_i} \cdot [u_\iota(\tilde{\mu}) - u_\iota(\hat{\mu})]. \tag{B.16}$$

As the local tree has symmetric losses it follows that $u_\iota(\tilde{\mu}) - u_\iota(\hat{\mu}) = u_{\iota'}(\tilde{\mu}) - u_{\iota'}(\hat{\mu})$; this follows from the fact that they both gain an equal number of new shortest paths through j, j', this follows as as j, j' have same number of paths after deletion of jj' due to symmetric losses. This entails that without loss of generality we can assume that $p_{i\iota} = \hat{p}_i$ as otherwise we could substitute ι with ι' and conduct the analysis again.

For ι and some agent i' of type $\tilde{x}$ it holds that $p_{ii'}(\tilde{\mu}) \leq p_{\iota i'}(\tilde{\mu}) + \hat{p}_i$. This follows as there exists a path between i, ι and ι, i' with respectively lengths $p_{\iota i'}(\tilde{\mu})$ and $\hat{p}_i$; thus $p_{ii'}(\tilde{\mu}) \leq p_{\iota i'}(\tilde{\mu}) + \hat{p}_i$. This implies the following inequality must hold:

$$\sum_{x_{i'}=\tilde{x}} \delta^{p_{ii'}(\tilde{\mu})} \geq \delta^{p_{\iota i}(\tilde{\mu})} \cdot \sum_{x_{i'}=\tilde{x}} \delta^{p_{\iota i'}(\tilde{\mu})}.$$

As $u_\iota(\tilde{\mu}) - u_\iota(\hat{\mu}) = \sum_{x_i=\tilde{x}} \prod_{l=1}^{p_{\iota i'}(\tilde{\mu})} \delta^{r_l} \cdot z(x, \tilde{x})$ and $u_i(\tilde{\mu}) - u_i(\hat{\mu}) = \sum_{x_i=\tilde{x}} \prod_{l=1}^{p_{ii'}(\tilde{\mu})} \delta^{r_l} \cdot z(x, \tilde{x})$ it follows that Inequality B.16 holds which proves our fact.

□

References

G. S. Becker, A Theory of Marriage: Part I, Journal of Political Economy 81 (1973) 813–846.

A. Bjerre-Nielsen, Sorting in networks: Adversity and structure, arXiv preprint arXiv:1503.07389 (2015).

M. O. Jackson, The Human Network: How Your Social Position Determines Your Power, Beliefs, and Behaviors, Pantheon Books, United States of America, 2019.

M. McPherson, L. Smith-Lovin, J. M. Cook, Birds of a Feather: Homophily in Social Networks, Annual Review of Sociology 27 (2001) 415–444.

M. O. Jackson, A. Wolinsky, A Strategic Model of Social and Economic Networks, Journal of Economic Theory 71 (1996) 44–74.

J. Ugander, B. Karrer, L. Backstrom, C. Marlow, The anatomy of the facebook social graph, arXiv preprint arXiv:1111.4503 (2011).

G. Miritello, E. Moro, R. Lara, R. Martinez-Lopez, J. Belchamber, S. G. Roberts, R. I. Dunbar, Time as a limited resource: Communication strategy in mobile phone networks, Social Networks 35 (2013) 89–95.

J. Church, I. King, Bilingualism and network externalities, Canadian Journal of Economics (1993) 337–345.

M. Kremer, The O-ring theory of economic development, The Quarterly Journal of Economics 108 (1993) 551–575.

P. S. Park, J. E. Blumenstock, M. W. Macy, The strength of long-range ties in population-scale social networks, Science 362 (2018) 1410–1413.

T. C. Schelling, Models of segregation, The American Economic Review 59 (1969) 488–493.

T. C. Schelling, Dynamic models of segregation, Journal of Mathematical Sociology 1 (1971) 143–186.

R. Shimer, L. Smith, Assortative matching and search, Econometrica 68 (2000) 343–369.

H. Chade, J. Eeckhout, L. Smith, Sorting through search and matching models in economics, Journal of Economic Literature 55 (2017) 493–544.

J. Farrell, S. Scotchmer, Partnerships, The Quarterly Journal of Economics 103 (1988) 279–297.

S. N. Durlauf, A. Seshadri, Is assortative matching efficient?, Economic Theory 21 (2003) 475–493.

P. Legros, A. F. Newman, Assortative Matching in a Non-Transferable World, SSRN Scholarly Paper ID 328460, Social Science Research Network, Rochester, NY, 2002.

M. Pycia, Stability and preference alignment in matching and coalition formation, Econometrica 80 (2012) 323–362.

M. Baccara, L. Yariv, Homophily in peer groups, American Economic Journal: Microeconomics 5 (2013) 69–96.

Y. Xing, Essays in microeconomic theory and the economics of networks, Ph.D. thesis, Stanford University, 2016. Chapter 1: Who Shares Risk with Whom and How? Endogenous Matching and Selection of Risk Sharing Equilibria.

J. M. Buchanan, An economic theory of clubs, Economica 32 (1965) 1–14.

J. de Martí, Y. Zenou, Segregation in friendship networks, The Scandinavian Journal of Economics 119 (2017) 656–708.

F. Bloch, M. O. Jackson, The formation of networks with transfers among players, Journal of Economic Theory 133 (2007) 83–110.

T. C. Koopmans, M. Beckmann, Assignment problems and the location of economic activities, Econometrica 25 (1957) 53–76.

C. Johnson, R. P. Gilles, Spatial social networks., Review of Economic Design 5 (2000) 273–299.

M. O. Jackson, B. W. Rogers, The economics of small worlds, Journal of the European Economic Association 3 (2005) 617–627.

A. Galeotti, S. Goyal, J. Kamphorst, Network formation with heterogeneous players, Games and Economic Behavior 54 (2006) 353–372.

V. Bala, S. Goyal, A noncooperative model of network formation, Econometrica 68 (2000) 1181–1229.

S. Currarini, M. O. Jackson, P. Pin, An economic model of friendship: Homophily, minorities, and segregation, Econometrica 77 (2009) 1003–1045.

S. Currarini, M. O. Jackson, P. Pin, Identifying the roles of race-based choice and chance in high school friendship network formation, Proceedings of the National Academy of Sciences 107 (2010) 4857–4861.

Y. Bramoullé, S. Currarini, M. O. Jackson, P. Pin, B. W. Rogers, Homophily and long-run integration in social networks, Journal of Economic Theory 147 (2012) 1754–1786.

J. D. Montgomery, Social networks and labor-market outcomes: Toward an economic analysis, The American Economic Review 81 (1991) 1408–1418.

M. Galenianos, Referral networks and inequality, Available at SSRN 2768083 (2018).

S. Goyal, F. Vega-Redondo, Structural holes in social networks, Journal of Economic Theory 137 (2007) 460–492.

P. Legros, A. F. Newman, Beauty Is a Beast, Frog Is a Prince: Assortative Matching with Nontransferabilities, Econometrica 75 (2007) 1073–1102.

A. S. J. Kelso, V. P. Crawford, Job matching, coalition formation, and gross substitutes, Econometrica 50 (1982) 1483–1504.

M. S.-Y. Chwe, Farsighted coalitional stability, Journal of Economic Theory 63 (1994) 299–325.

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

P. J.-J. Herings, A. Mauleon, V. Vannetelbosch, Farsightedly stable networks, Games and Economic Behavior 67 (2009) 526–541.

Y. Bachrach, E. Elkind, R. Meir, D. Pasechnik, M. Zuckerman, J. Rothe, J. S. Rosenschein, The cost of stability in coalitional games, in: M. Mavronicolas, V. G. Papadopoulou (Eds.), Algorithmic Game Theory, Springer, Berlin, Heidelberg, 2009, pp. 122–134.

B. Klaus, M. Walzl, Stable many-to-many matchings with contracts, Journal of Mathematical Economics 45 (2009) 422–434.