Rev. Econ. Design (2012) 16:71–87
DOI 10.1007/s10058-012-0114-x

ORIGINAL PAPER

Under-connected and over-connected networks: the role of externalities in strategic network formation

Berno Buechel · Tim Hellmann

Received: 17 February 2011 / Accepted: 31 January 2012 / Published online: 15 February 2012

Abstract Since the seminal contribution of Jackson and Wolinsky (J Econ Theory 71(1):44–74, 1996) it has been widely acknowledged that the formation of social networks exhibits a general conflict between individual strategic behavior and collective outcome. What has not been studied systematically are the sources of inefficiency. We approach this omission by analyzing the role of positive and negative externalities of link formation. This yields general results that relate situations of positive externalities with stable networks that cannot be "too dense" in a well-defined sense, while situations with negative externalities tend to induce "too dense" networks. Those results are neither restricted to specific assumptions on the agents' preferences (e.g. homogeneity), nor to a specific notion of stability or efficiency.

Keywords Networks · Connections · Externalities · Spillovers · Stability · Efficiency

JEL Classification D85 · C72 · L14

1 Introduction

As in non-cooperative game theory, a central issue in the theory of network formation is the analysis of equilibrium or stability, i.e. a situation where no player wants to change

We thank Herbert Dawid, Bhaskar Dutta, Matt Elliot, Matt Jackson, Jurjen Kamphorst, Paolo Pin, Frank Riedel, Walter Trockel, and the audiences from the conferences CTN (Maastricht), ESEM (Barcelona), ICFSN (Paris), SAET (Ischia) and SED (Maastricht) for helpful comments and suggestions.

B. Buechel (✉)
Department of Economics, University of Hamburg, Hamburg, Germany
e-mail: berno.buechel@wiso.uni-hamburg.de

T. Hellmann
Institute of Mathematical Economics, Bielefeld University, Bielefeld, Germany
e-mail: thellmann@wiwi.uni-bielefeld.de

his links. Myerson (1991) proposes a simultaneous move game of network formation, where players announce their desired links non-cooperatively, but the standard formulation of Nash equilibrium in the Myerson network formation game has proved to be a non-satisfying concept due to coordination problems.[1] In the seminal contribution of Jackson and Wolinsky (1996) this problem is solved by introducing a different concept of stability called pairwise stability. In a pairwise stable network no two players want to form a mutual link and no player wants to cut a link unilaterally. This concept of stability is used and has been refined widely in the literature of network formation games. Jackson and Wolinsky (1996) highlight a central problem in strategic network formation: There is a tension between stability and efficiency, meaning that individual interest can be at odds with societal welfare.[2] Since then, there was a flourishing literature on specific situations of strategic network formation of which two small surveys can be found in Jackson (2004) and Goyal and Joshi (2006b). Indeed, in various models it can be observed that stable networks do not coincide with efficient networks.

What has not been explicitly studied are the *sources of inefficiency*. The question is particularly, how do stable networks generally differ from efficient networks? And, being in an inefficient (e.g. stable) situation, how can welfare be improved without reshuffling the whole network structure?

We approach these questions by analyzing the role of externalities (also called spillovers) of link formation. Simply put, positive externalities define situations where agents can profit (at least do not suffer) from others who form a relationship; while negative externalities mean that they do not benefit from that action. We argue that both types of externalities correspond to natural settings. Network formation games where direct and indirect connections are the source of benefits represent examples for positive externalities. On the other hand, in a context of competition or rival goods, negative externalities occur.

For our analysis, we employ several notions of stability and efficiency. In particular, we use three well-known notions of stability: pairwise stability, as introduced in Jackson and Wolinsky (1996), pairwise Nash stability, a simple refinement of the former which incorporates the property of Nash equilibrium, and pairwise stability with transfers, which stems from a network formation game allowing for transfers.[3] For the analysis of the welfare properties of the stable networks, we use a very general set of welfare functions, which have to satisfy only a monotonicity property.[4] Given a welfare function, we introduce the notion of *over-connected* and *under-connected* networks. While usually networks are classified as either efficient (welfare maximizing) or

[1] Any link that is desired by both players is not necessarily present in Nash equilibrium if neither player announces it, e.g. the empty network is always an equilibrium.

[2] See Jackson and Wolinsky (1996), Theorem 1, for the general statement about the tension of pairwise stable and efficient networks. Note that this can be relaxed for strongly stable networks (see Dutta and Mutuswami 1997, Theorem 4.19.).

[3] See Bloch and Jackson (2007) for different approaches to network formation with transfers. For a comparison of the equilibrium concepts see Bloch and Jackson (2006) and Calvó-Armengol and Ilkiliç (2009).

[4] The utilitarian welfare function, the sum of individual utilities, satisfies this notion. For some of the results, we actually need this specific version of a welfare function.

inefficient, the two notions further describe inefficient networks. In essence, a network is over-connected if welfare can be improved by deleting some links, while a network is under-connected if an addition of links is welfare improving. We show how these notions help identify the sources of inefficiency and can be applied to characterizing stable and efficient networks.

The main result for positive externalities is that there is no stable network that is over-connected (Theorem 1). This result holds directly for the notion of pairwise Nash stability, while it takes an additional assumption ("concavity") to make it hold for the notion of pairwise stability. The result is, however, not dependent on the particular shape of the utility functions nor on the degree of homogeneity. A direct interpretation of Theorem 1 is that under positive externalities a stable network cannot be socially improved by the severance of links. In fact, the statement is even stronger: Under positive externalities any network that is contained in a pairwise Nash stable network is weakly Pareto dominated by the former. We illustrate the implications of the result in an example taken from the literature, the *connections model* (cf. Jackson and Wolinsky 1996).

For negative externalities a corresponding result cannot be established immediately. In the context of transfers, however, there is an analogous result: No stable network is under-connected (Theorem 2). This means that under negative externalities no stable network can be socially improved by the addition of links. While the qualitative statement is analogue to Theorem 1 on positive externalities, the two results differ significantly: Theorem 2 applies to the notion of pairwise stability with transfers, it requires an additional assumption ("concavity"), and it is restricted to a special case of a welfare function (the sum of individual utilities). We discuss why these restrictions are needed under negative externalities and how they can be relaxed. Finally, we apply Theorem 2 to a model of *patent races* (Goyal and Joshi 2006b) to illustrate how our results can contribute to and extend previous results on the characterization of the set of stable networks (Proposition 2).

Our results are applicable to many network formation games from the literature. Examples for positive externalities include the *provision of a pure public good* and a model of *market sharing agreements* (both introduced in Goyal and Joshi 2006b); and the *connections model* (Jackson and Wolinsky 1996). Examples for negative externalities include the model of *patent races* (Goyal and Joshi 2006b), a *co-author model* (Jackson and Wolinsky 1996), and a model of *free trade agreements* (Goyal and Joshi 2006a).

This paper is organized as follows: The subsequent Sect. 2 formally defines the model. The results on positive externalities are presented and discussed in Sect. 3. The results on negative externalities are presented and discussed in Sect. 4. In Sect. 5 we conclude.

2 Model and definitions

Let $N = \{1, \ldots, n\}$ be a set of agents/players, with $n \geq 3$. A network g is a set of unordered pairs $\{i, j\}$ with $i \neq j \in N$, that represent the bilateral relations. Thus, $ij := \{i, j\} \in g$ means that player i and player j are linked in network g.

Let g^N be the set of all subsets of N of size two and let G be the set of all possible graphs, $G = \{g : g \subseteq g^N\}$. By $N_i(g)$ we denote the neighbors of player i in network g, $N_i(g) := \{j \in N \mid ij \in g\}$. Similarly, $L_i(g)$ denotes the set of player i's links in g, $L_i(g) := \{ij \in g \mid j \in N\}$. We define $d_i(g) := |L_i(g)| = |N_i(g)|$, as the number of player i's links, called player i's degree.

For each player $i \in N$ a utility function $u_i : G \to \mathbb{R}$ expresses his preferences over the set of possible graphs. $u = (u_1, \ldots, u_n)$ denotes the profile of utility functions. Decisions to form or to sever links typically do not depend on absolute utility, but on changes in utility. Let $mu_i(g, l)$ be the marginal utility of player i of deleting a set of links l in network g, that is $mu_i(g, l) := u_i(g) - u_i(g \backslash l)$ for $l \subseteq g$. Equivalently, we denote $mu_i(g \cup l, l) := u_i(g \cup l) - u_i(g)$ as the marginal utility of adding the set of links l to network g.

From the vast literature of network formation, we employ three of the most common stability notions. The first notion is based on a cooperative framework and was introduced by Jackson and Wolinsky (1996).

Pairwise Stability. A network g is pairwise stable (PS) if no link will be cut by a single player, and no two players want to form a link:

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

This well-known definition captures the idea that links can be severed by any involved player, whereas the formation of a link requires the consent of both players. Pairwise stability is a basic notion that can be refined in multiple ways. One of the refinements is the property of pairwise Nash stability. A network is pairwise Nash stable (PNS) if there exists a Nash equilibrium in the corresponding link formation game (see Myerson 1991) that supports this network and no link will be added by two players. This boils down to the following conditions.

Pairwise Nash Stability. A network g is pairwise Nash stable (PNS) if the following holds:

(i) $\forall i \in N, \quad \nexists l \subseteq L_i(g) : u_i(g \backslash l) > u_i(g)$ and
(ii) $\forall ij \notin g, \quad u_i(g \cup ij) > u_i(g) \Rightarrow u_j(g \cup ij) < u_j(g)$.

In contrast to pairwise stability, pairwise Nash stability captures the idea that players are able to delete multiple links simultaneously. The third notion of stability is based on the idea of transfers and can be found in Bloch and Jackson (2007).

Pairwise Stability with Transfers. A network g is pairwise stable with transfers (PST) if there does not exist any pair of players that can jointly benefit by adding, respectively cutting, their link:

(i) $\forall ij \in g, \quad u_i(g) + u_j(g) \geq u_i(g \backslash ij) + u_j(g \backslash ij)$ and
(ii) $\forall ij \notin g, \quad u_i(g) + u_j(g) \geq u_i(g \cup ij) + u_j(g \cup ij)$.

We denote by $[PS(u)]$, $[PNS(u)]$, and $[PST(u)]$ the sets of pairwise stable, pairwise Nash stable, and pairwise stable networks with transfers, respectively, for a given profile u of utility functions.

While stability tries to answer which networks emerge based on individual preferences, efficiency addresses the evaluation of networks from a societal point of view.

To formally capture efficiency, we use a welfare function $w : G \to \mathbb{R}$ that typically (but not necessarily) is only dependent on the vector of utilities of all players, given a network g. The most commonly used version of a welfare function is the utilitarian welfare function, which simply sums up the utility of all players, $w^u(g) = \sum_{i \in N} u_i(g)$. For some of our results, however, an even weaker way of aggregating utility is sufficient. We only require a welfare function to satisfy the following property.

Definition 1 A welfare function w satisfies monotonicity if

(i) $u_i(g) \geq u_i(g') \;\; \forall i \in N \implies w(g) \geq w(g')$, and
(ii) $u_i(g) \geq u_i(g') \;\; \forall i \in N$ and $\exists j \in N : u_j(g) > u_j(g') \implies w(g) > w(g')$.

This assumption is a very weak requirement for a welfare function: A welfare function should evaluate a network g at least as high as a network g' if all players $i \in N$ evaluate g at least as high as g'. Monotonicity assures that the welfare function preserves the Pareto ordering of the networks. Given a welfare function w, let us define efficiency.

Efficiency. A network g^* is called efficient with respect to the welfare function w if it is a welfare maximizing network, that is $w(g^*) \geq w(g) \;\; \forall g \in G$.

We introduce the following two definitions in order to describe non-efficient networks.

Definition 2 A network g is called *over-connected* (with respect to the welfare function w) if $\exists g' \subset g$ such that $w(g') > w(g)$.

Definition 3 A network g is called *under-connected* (with respect to the welfare function w) if $\exists g' \supset g$ such that $w(g') > w(g)$.

A network is over-connected if it is "too dense" in the sense that overall welfare can be improved by cutting links. Similarly, under-connected networks are "not dense enough". Efficient networks are neither over-connected nor under-connected. Inefficient networks, however, can satisfy both, one, or none of these two properties. To shed some light into the tension between stability and efficiency, we will ask whether and under what conditions stable networks are over-connected or not under-connected, respectively under-connected or not over-connected.

3 Positive externalities

Positive externalities in network formation games simply capture that players experience positive effects on their utility when others form a link. As defined below, a link formed by two players cannot decrease other players' utility.[5]

Definition 4 A utility function u_i satisfies positive externalities if $\forall g \in G, \forall jk : i \notin \{j, k\}$, it holds that

$$u_i(g \cup jk) \geq u_i(g).$$

[5] Externalities in this case capture the effects of the decision of two players forming a link on other players' utility. This is different to the meaning of externalities in the context of markets. Note also that we have not required that the inequalities are strict.

A profile of utility functions u satisfies positive externalities if all utility functions satisfy positive externalities.

Being required for any network, any link, and any player, this property seems to be quite restrictive. However, we argue that this property is apparent in many contexts of network formation. Some examples from the literature are *provision of a pure public good* (Goyal and Joshi 2006b), *market sharing agreements* (Belleflamme and Bloch 2004), and the *connections model* (Jackson and Wolinsky 1996), which we discuss below.

In case of a utility function that is additive separable into costs and benefits, positive externalities are implied by a simple monotonicity property of the benefit function. In this context, players have to carry the costs of their own links, but share the benefits with others. Intuitively, individual incentives to establish a link might be lower than its collective value because of positive externalities.

Theorem 1 formalizes this intuition. More precisely, it shows that a pairwise Nash stable network can never be socially improved by the deletion of links.

Theorem 1 *If a profile of utility functions u satisfies positive externalities, then no pairwise Nash stable network is over-connected with respect to any monotonic welfare function w.*

The proof of this and all following statements can be found in the "Appendix". To prove this result, we show that no player is better off in a subnetwork g' of a pairwise Nash stable network g. Pairwise Nash stability implies that a player cannot prefer a network $\tilde{g}(\subset g)$ that has only been reduced by some of his own links. Because of positive externalities, he cannot prefer a subnetwork $g' \subset \tilde{g}$ of the reduced network. The argument holds for any player such that the monotonicity property of the welfare function establishes the result. Note that this also implies that a pairwise stable network weakly Pareto dominates *any* subnetwork: Deleting links makes no player better off.

In the proof of Theorem 1 we use the property of pairwise Nash stable networks that no player can benefit by unilaterally severing a set of own links.[6] With the notion of pairwise stability, however, each link is considered one by one. In order to apply Theorem 1 to the weaker notion of pairwise stability we need an additional assumption.

Definition 5 A utility function u_i is concave (in own links) if $\forall g \in G$, and $\forall l_i \subset L_i(g^N \backslash g)$, $\forall ij \in g$ it holds that

$$mu_i(g, ij) \geq mu_i(g \cup l_i, ij).$$

A profile of utility functions is concave if all utility functions are concave.

The property requires that the marginal utility of a link is decreasing in the set of links the player has already. Hellmann (2009) shows that this concavity property is equivalent to two notions we find in the literature (see Lemma 2 in the "Appendix").

6 Note that we picked pairwise Nash stability for illustrative reasons. In fact the proof of Theorem 1 shows that Nash stability, i.e. property (i) of pairwise Nash stability, is already sufficient for the result.

Concavity builds a bridge between the notion of pairwise Nash stability and pairwise stability: If a profile of utility functions is concave, then the two stability notions coincide, that is $[PNS(u)] = [PS(u)]$.[7] Thus, we get the following Corollary to Theorem 1.

Corollary 1 *Suppose a profile of utility functions u satisfies positive externalities and concavity, then no pairwise stable network is over-connected with respect to any monotonic welfare function w.*

What has been implied by pairwise Nash stability in Theorem 1, is now assured by pairwise stability together with the assumption of concavity: In a stable network no player can improve by cutting a set of his links. The notions of stability that are used in Theorem 1 and Corollary 1 are fairly weak. The statement, also applies to any refinement of pairwise stability, e.g. bilateral stability (Goyal and Vega-Redondo 2007), unilateral stability (Buskens and van de Rijt 2008), strong stability (Dutta and Mutuswami 1997), and weak stability (Dutta and Mutuswami 1997). Moroever, for the notion of pairwise stability with transfers it is possible to state a similar result, which is however restricted to the utilitarian welfare function.

Theorem 1 excluding over-connectedness has trivial implications for the complete and empty network: As any network is a subnetwork of the complete network, it follows that (a) if the complete network is stable, then it must also be efficient. Since any network is a supernetwork of the empty network it follows that, (b) if the empty network is uniquely efficient, then no other network can be stable. Next, we show how Theorem 1 applies to a model from the literature.

The connections model revisited

The connections model was introduced in Jackson and Wolinsky (1996). It models the flow of resources (like information or support) via shortest paths in a network. Let $d_{ij}(g)$ denote the distance of players i and j in network g (which is defined to be ∞ for unconnected pairs), then the utility of each player can be written as

$$u_i^{CO}(g) = w_{ii} + \sum_{j \neq i} \delta^{d_{ij}(g)} w_{ij} - \sum_{j: ij \in g} c_{ij}, \quad with \;\; \delta \in (0, 1). \tag{1}$$

w_{ij} stands for the undiscounted value of a connection to player j and c_{ij} stands for the cost of maintaining a link with agent j. It is easy to see that the connections model satisfies positive externalities. If ij forms in some network g, then the utility of a player $k \neq \{i, j\}$ either does not change or increases as some of k's distances are shortened because $d_{km}(g \cup ij) \leq d_{km}(g)$ for all m. Consequently (by Theorem 1), no Nash stable network can be over-connected w.r.t. any monotonic welfare function.

Moreover, it can be shown that $u^{CO}(\cdot)$ satisfies concavity. By the result of Hellmann (2009) it suffices to show that u^{CO} satisfies convexity in own current links, that is

[7] See Calvó-Armengol and Ilkiliç (2009).

$\forall i \in N, \forall g \in G$, and $\forall l \subseteq L_i(g)$, it holds that $mu_i(g,l) \geq \sum_{ij \in l} mu_i(g,ij)$. This has been done by Calvó-Armengol and Ilkiliç (2009) for the symmetric connections model. We make a straightforward generalization of their proof.[8]

Lemma 1 *The heterogeneous connections model satisfies concavity in own links.*

Consequently (by Corollary 1), no pairwise stable network can be over-connected w.r.t. any monotonic welfare function. While stable networks depend on the dyadic specifications of value and costs (w_{ij}, c_{ij}), the results excluding over-connectedness imply that the welfare of a stable network can never be improved by severing links.

There are more specific results for the connections model in its symmetric version, setting $w_{ij} = 1, c_{ij} = c(\forall i \neq j)$ and considering the utilitarian welfare function w^u only. This has been studied in Jackson and Wolinsky (1996), Jackson (2003), and Hummon (2000). Jackson and Wolinsky (1996, Prop. 1 and Prop. 2) show that for low costs ($c < \delta - \delta^2$) the complete network is efficient (and uniquely pairwise stable); for medium costs ($\delta - \delta^2 < c < \delta + \frac{n-2}{2}\delta^2$) the star network is efficient; while for very high costs ($c > \delta + \frac{n-2}{2}\delta^2$) the empty network is efficient. The famous statement of inefficiency in the connections model is the following: "For $\delta < c$, any pairwise stable network which is non-empty is such that each player has at least two links and thus is inefficient".[9]

What does our result excluding over-connectedness add to this discussion of inefficiency? First, there is the above mentioned trivial implication (b) for the empty network: Since any network is a supernetwork of the empty network, it follows that if the empty network is uniquely efficient, then no other network can be stable. Thus, the statement of inefficiency is restricted to $\delta < c < \delta + \frac{n-2}{2}\delta^2$. Second, the result on over-connectedness adds a new point of view on the flavor of inefficiency. This can be illustrated in the following example which is also taken from Jackson and Wolinsky (1996, Ex. 1).

Example 1 The network in Fig. 1, called "Tetrahedron", is stable for costs $c > \delta$, where the star network is uniquely efficient.[10] The Tetrahedron is "too dense" in the sense that it has 18 links, while the efficient network has 15. Accordingly, Jackson and Wolinsky (1996, p. 51) label it as "over-connected". However, by Theorem 1 and Corollary 1, it is not over-connected according to the definition used in this paper. This means that the welfare of the Tetrahedron cannot be improved by leaving out any set of its links. Moreover, we claim that the Tetrahedron is under-connected in the parameter range for which it is pairwise stable. In the "Appendix" (Proposition 1) we show that the addition of a link between the players "2" and "13" would strictly improve utilitarian welfare. The same point as in the Tetrahedron can be illustrated in many other stable networks in the connections model: they are under-connected for any costs for which they are pairwise stable.

[8] In fact, such a generalization can be made for any "distance-based" utility function in the sense that benefits are decreasing with distances and costs only depend on direct links.

[9] cf. Jackson and Wolinsky (1996), p. 51.

[10] More precisely, g^{Tetra} is pairwise stable iff $\delta - \delta^5 + \delta^2 - \delta^4 + \delta^2 - \delta^5 + 2(\delta^3 - \delta^4) \leq c \leq \delta - \delta^8 + \delta^2 - \delta^7 + \delta^3 - \delta^6 + 2(\delta^4 - \delta^5)$.

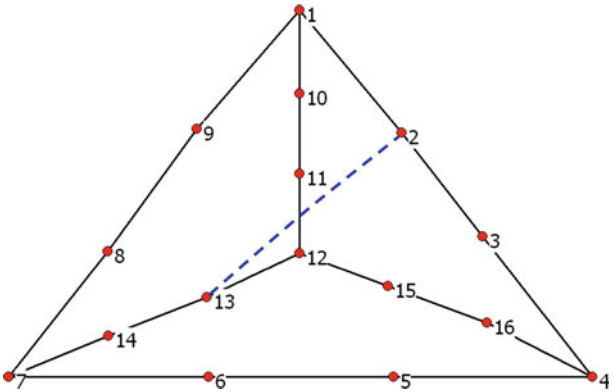

Fig. 1 Example of an inefficient network ("Tetrahedron")

The example illustrates two different viewpoints on inefficiency. From the viewpoint of a social planner that can unrestrictedly manipulate a given network, some stable networks are "too dense" in the sense that less links are needed to form the efficient one. From the viewpoint of a social planner who is restricted to either foster or hinder the formation of links (e.g. by taxes or subsidies), many stable networks in the connections model are "not dense enough" (under-connected), while none is "too dense" (over-connected).

4 Negative externalities

Negative externalities in network formation occur, when the addition of a link cannot be beneficial for the players who are not involved in this link. Formally, we speak of negative externalities if the following holds.

Definition 6 A utility function u_i satisfies negative externalities if $\forall g \in G, \forall jk : i \notin \{j, k\}$, it holds that

$$u_i(g \cup jk) \leq u_i(g).$$

A profile of utility functions u satisfies negative externalities if all utility functions satisfy negative externalities.

When considering negative externalities in economics, equilibrium analysis usually shows that individuals rather do "too much" (pollute, etc.) than being socially optimal. Players facing negative externalities are thus expected to create "too dense" networks. One example is the *co-author model* introduced in Jackson and Wolinsky (1996). In this model, a researcher who forms a new collaboration does not internalize the negative effects on her current collaborators. Formally, it can be shown that any pairwise stable network in this *co-author model* is over-connected with respect to utilitarian welfare. However, such a result does not hold in all models satisfying negative externalities. It may even happen that a stable network is under-connected because of the following reason. It might happen that a player rejects a link that is highly beneficial to the proposing player. This link could potentially lead to higher welfare.[11]

A simple way out of this issue is using the stability concept pairwise stability with transfers. This concept helps ensure that any single link that is not contained in a pairwise stable network with transfers cannot be welfare improving. Analogously to Corollary 1, we require concavity (see Definition 5) for our main result on negative externalities to ensure that there is no set of links that would improve welfare.

Theorem 2 *Suppose a profile of utility functions u satisfies negative externalities and concavity, then no network $g \in G$, which is pairwise stable with transfers, is under-connected with respect to the utilitarian welfare function.*

As always, the proof can be found in the "Appendix". Pairwise stability with transfers differs from pairwise stability significantly. In general, neither $[PS(u)] \subseteq [PST(u)]$ nor $[PS(u)] \supseteq [PST(u)]$. Although the concepts differ, we can easily find properties of the utility function that are sufficient to ensure equivalence of both stability concepts. It can be shown that the property of pairwise sign compatibility in Chakrabarti and Gilles (2007) is sufficient for $[PS(u)] = [PST(u)]$. Thus, from Theorem 2 we can conclude that also no pairwise stable network (and hence no Pairwise Nash stable network) is under-connected in the presence of pairwise sign compatibility, negative externalities, and concavity.

For the results, so far, it has been crucial whether network formation requires bilateral consent or can be done unilaterally. In the definitions of pairwise stability and pairwise Nash stability, it is implicitly assumed that link deletion can be done unilaterally, while link creation requires the consent of both involved players. If we allow for opposite link formation rules, then the assumptions made for the result on negative externalities, i.e. Theorem 2, are required for excluding over-connectedness in the presence of positive externalities, while the non-under-connectedness of stable networks in case of negative externalities are then analogous to the results we have on positive externalities.[12] For example, pairwise Nash stability incorporates deviations of simultaneously deleting several links. If we consider a stability notion that incorporates the players' ability to unilaterally *add* a set of links, we can formulate a result that is fully analogous to Theorem 1: Networks that are stable in that sense, can be shown to be not under-connected with respect to any monotonic welfare function.

There are many examples in the literature of network formation which satisfy negative externalities (and other properties required for Theorem 2).[13] The *co-author* model introduced in Jackson and Wolinsky (1996) is mentioned above. Goyal and Joshi (2006a) present a model of *free trade agreements* that satisfies negative externalities. In order to apply Theorem 2 we need concavity, which only holds on a restricted domain, $\tilde{G} := \{g \in G \mid \ d_i(g) \geq 1 \ \ \forall i \in N\}$. However, we can extend Theorem 2

[11] This issue can also be considered as some kind of external effect. However, we prefer to distinguish utility considerations of the directly involved agents from effects to non-involved agents.

[12] Thus, our results can be restated in terms of bilateral or unilateral link creation/deletion rules or by using notions of addition proofness and deletion proofness.

[13] Any example mentioned in this paper is discussed more extensively (with respect to our results) in an earlier working paper version (Buechel and Hellmann 2009).

straightforwardly to a restricted domain. Thus, we conclude that no stable network $g \in \tilde{G}$ is under-connected within the set of networks $\tilde{G}$ in that model of free trade agreements. Finally, Goyal and Joshi (2006b) introduce a model of Patent Races. As an example of Theorem 2, we discuss this model more intensively in the next subsection.

Patent races

Goyal and Joshi (2006b) derive this model as a variation of the classical patent race model.[14] In addition to the classical model, firms can join R&D collaborations to accelerate research. The first firm to develop the new product is awarded a patent. The random time $\tau(d_i(g))$ at which the innovation happens is given by

$$Pr(\{\tau(d_i(g)) \leq t\}) = 1 - \exp(-d_i(g)t).$$

Assuming risk neutrality, payoff of 1 in case of receiving the patent and 0 else, and a discount factor ρ, the expected payoff of firm i is the following:

$$\begin{aligned} & u_i^{PR}\big(d_i(g), D(g_{-i})\big) \\ & = E_t\Big[exp(-\rho t)Pr(\tau(d_i(g)) = t)\prod_{j\neq i} Pr(\tau(d_j(g)) > t)\Big] - d_i(g)c \\ & = \frac{d_i(g)}{\rho + D(g)} - d_i(g)c = \frac{d_i(g)}{\rho + 2d_i(g) + D(g_{-i})} - d_i(g)c, \end{aligned}$$

where g_{-i} represents the network obtained by deleting player i and all his links and $D(g) := \sum_{i\in N} d_i(g)$. This model satisfies negative externalities since links of other firms reduce the probability to innovate first. Also, since u_i^{PR} is a concave function of $d_i(g)$, it is concave according to Definition 5. From Theorem 2 we can thus conclude that no pairwise stable network with transfers is under-connected. In fact, it is straightforward to calculate the efficient networks in this model since the utilitarian welfare is given by:

$$w^{PR}(g) = \sum_{i\in N} u_i^{PR}(g) = \sum_{i\in N}\left(\frac{d_i(g)}{\rho + D(g)} - d_i(g)c\right) = \frac{D(g)}{\rho + D(g)} - D(g)c.$$

In this case the utilitarian welfare only depends on the total number of links and thus any network that contains the optimal number of total links is efficient. The distribution of links and the structure of the network do not matter for efficiency. We can easily calculate that for $\frac{\rho}{(\rho+2(k+1))(\rho+2k)} < c < \frac{\rho}{(\rho+2k)(\rho+2(k-1))}$ any network which contains k links is efficient and no other networks are efficient.

It requires a little bit more to characterize stable networks. However, for this matter we can apply Theorem 2 in order to bound the total number of links.

Proposition 2 *Suppose that* $\frac{\rho}{(\rho+2k+2)(\rho+2k)} < c$*, then all networks* g *which are pairwise stable with transfers have to contain more than* k *links, in other words* $D(g) \geq 2k$.

[14] See Dasgupta and Stiglitz (1980) among others.

In their paper, Goyal and Joshi (2006b) only find a partial characterization for the set of pairwise stable networks. By applying Theorem 2 we were able to contribute to their characterization. This example shows that our theorems not only describe the tension between stability and efficiency, but can also be applied to characterize the stable networks (resp. the efficient ones).

5 Concluding remarks

We have introduced the notion of over-connected and under-connected networks in order to contribute to a better understanding of the tension between stability and efficiency in situations of strategic network formation. An over-connected network can be socially improved by the deletion of links; an under-connected network can be socially improved by the addition of links. In that way we relate inefficient outcomes to externalities of link formation.

The basic argument is that positive spillovers/externalities lead to situations where agents are not willing to form links, although it would be collectively beneficial. Negative externalities have the opposite effect: agents form links without internalizing the loss of utility of other agents. It is illustrated in specific models, that inefficient networks in one setting (positive externalities) are under-connected (e.g. the connections model), while the inefficient networks in the other setting (negative externalities) are over-connected (e.g. the co-author model). However, this observation does not hold in general. What can be shown generally, is the following: For positive externalities no stable network can be over-connected (Theorem 1 and Corollary 1); for negative externalities no stable network can be under-connected when some other conditions are met (Theorem 2). Thus, the stable networks in one setting cannot be improved by the deletion of links, while the stable networks in the other setting cannot be improved by the severance of links.

Despite their intuitive character, our results are not trivial. As the analysis shows externalities are not the only source of inefficiency. Other sources of inefficiencies in the bilateral formation of links are miscoordination of actions in the Consent Game,[15] restrictions on possible deviations, and rejection power in the Consent Game. These three sources of inefficiency are addressed in the basic notions of stability that we use in our paper. Pairwise stability, for instance, solves the issue of miscoordination in the Consent Game. The second issue, however, is a problem of pairwise stability itself since links are considered one by one. A set of links can have contrary effects than each single link. This is solved by considering pairwise Nash stability (or by assuming concavity of the utility functions). The third case of inefficiency can be illustrated in the example where a player rejects a link although the partner would have benefited heavily from it. This can be ruled out by the introduction of transfers (as used in Theorem 2).

In the paper, we have restricted our attention to the standard formalization of networks as undirected graphs without weights. It can be shown that our results can be straightforwardly extended to the formation of directed unweighted and directed

[15] By *Consent Game* we mean the link announcement game introduced in Myerson (1991).

weighted networks as introduced in Bala and Goyal (2000), respectively Rogers (2006). Furthermore, for positive externalities it is possible to show that Nash equilibria are not over-connected in a framework of formation of undirected and weighted networks as introduced, e.g. in Bloch and Dutta (2009).[16]

The contribution of our results is two-fold. First, they shed light into the general tension between stability and efficiency giving a social planner a clear signal in which situations rather to impede and when to promote the formation of relationships. Second, the results can be used in specific models to improve the characterization of stable and efficient networks. We have illustrated this with two examples, while there are many other models of strategic network formation that meet the required conditions. In particular, our results do not rely on assumptions of homogeneous agents, nor on restrictions to certain notions of stability or efficiency. We hope that future research will come up with more of such interesting models accounting for the various nature of social and economic relationships.

Appendix

Proof of Theorem 1 Let $g \in [PNS(u)]$ and suppose that u satisfies positive externalities. We show that for all $g' \subset g$ it holds that $u_i(g') \leq u_i(g)$ for all $i \in N$. Let $l := l(g, g') = g \backslash g'$ for some $g' \subset g$, and denote $l_i := l_i(g, g') = l \cap L_i(g)$ and $l_{-i} := l \backslash l_i$. Since g is pairwise Nash stable, all owners of a link prefer to have all their links in g, i.e. $u_i(g) \geq u_i(g \backslash l_i)$.

Since u satisfies positive externalities, it holds for $\tilde{g} := g \backslash l_i$ that $u_i(\tilde{g}) \geq u_i(\tilde{g} \backslash l_{-i})$ (because player i does not own a link in l_{-i}), i.e. $l_{-i} \cap L_i(g) = \emptyset$). Therefore: $u_i(g) \geq u_i(g \backslash l_i) \geq u((g \backslash l_i) \backslash l_{-i}) = u(g')$. The same argument holds for all $i \in N$, implying that $w(g) \geq w(g')$ for any welfare function satisfying monotonicity. □

The following lemma from Hellmann (2009) is required for the proof of Corollary 1, Theorem 2, and Lemma 1. Recall the formal definitions of concavity and convexity from the literature. *Convexity in own current links* (Bloch and Jackson 2007): A profile of utility functions u is convex in own current links if $\forall i \in N, \forall g \in G$, and $\forall l \subseteq L_i(g)$ it holds that $mu_i(g, l) \geq \sum_{ij \in l} mu_i(g, ij)$. *(1-)Concavity in own* new *links* (Calvó-Armengol and Ilkiliç 2009): A profile of utility functions u is concave in own new links if for all $i \in N$, for all $g \in G$ and for all links l such that $l \subseteq L_i(g^N)$, and $l \cap g = \emptyset$ the following holds: $mu_i(g \cup l, l) \geq \sum_{ij \in l} mu_i(g \cup ij, ij)$.[17] For the definition of concavity in own links, see Definition 5.

Lemma 2 (Hellmann 2009)
The following statements are equivalent:

(1) *u is concave (convex) in own links.*

[16] In an earlier working paper version of this paper, we have introduced a general framework that allows for the formation of undirected and (possibly) weighted networks, as well as directed and (possibly) weighted networks. We show that the central results can be reestablished in this more general setup (see Buechel and Hellmann 2009).

[17] In the published version Calvó-Armengol and Ilkiliç (2009) relabel this property as 1-strong submodularity.

(2) *u is concave (convex) in own new links.*
(3) *u is convex (concave) in own current links.*

While it may seem odd that convexity in own current links is equivalent to concavity in own new links and concavity in own links, the curvature of the utility function is the same in all setups. Convexity in own current links is a somewhat unusual definition of convexity of a utility function, since the effects of *deleting* links need to have a convex curvature, which implies a concave curvature of the marginal effects of *adding* links. The formal proof of equivalence can be found in Hellmann (2009).

Proof of Corollary 1 By Lemma 2 above concavity in own links is equivalent to convexity in own current links. Calvó-Armengol and Ilkiliç (2009) show that (1-)convexity in own links is sufficient for $[PS(u)] = [PNS(u)]$. Thus, Theorem 1 applies. □

Proof of Theorem 2 Let g be pairwise stable with transfers. We show that for all $g' \supset g$ it holds that $\sum_{i\in N} u_i(g') \leq \sum_{i\in N} u_i(g)$. Suppose that u satisfies negative externalities and concavity. For $g' \supset g$, let $l := g'\backslash g$ and for each $i \in N$ let $l_i = l \cap L_i(g')$ and $l_{-i} := l\backslash l_i(g, g')$. Since u satisfies negative externalities, it holds for all $i \in N$ that:

$$u_i(g') \leq u_i(g'\backslash l_{-i}). \tag{2}$$

Concavity is equivalent to concavity in own new links, which implies for all $i \in N$:

$$u_i(g \cup l_i) - u_i(g) \leq \sum_{j:ij\in l_i} u_i(g \cup ij) - u_i(g). \tag{3}$$

Now, since g is pairwise stable with transfers (PST), (2) and (3) imply:

$$\begin{aligned}
\sum_{i\in N}(u_i(g') - u_i(g)) &= \sum_{i\in N}(u_i(g \cup l_i \cup l_{-i}) - u_i(g)) \\
&\overset{(2)}{\leq} \sum_{i\in N}(u_i(g \cup l_i) - u_i(g)) \\
&\overset{(3)}{\leq} \sum_{i\in N}\Big(\sum_{j:ij\in l_i}[u_i(g \cup ij) - u_i(g)]\Big) \\
&\overset{(*)}{=} \sum_{ij\in l} u_i(g \cup ij) - u_i(g) + u_j(g \cup ij) - u_j(g) \overset{(PST)}{\leq} 0,
\end{aligned}$$

where the equality (*) holds because for each link $ij \in l$ it holds that $ij \in l_1(k)$ if and only if $k \in \{i, j\}$ and only links in l are considered. □

Proof of Lemma 1 By Lemma 2 it suffices to show that u^{CO} satisfies convexity in own current links, that is $\forall i \in N, \forall g \in G$, and $\forall l \subseteq L_i(g)$, it holds that $mu_i^{CO}(g, l) \geq \sum_{ij\in l} mu_i^{CO}(g, ij)$.

Denote $\kappa_i(g, l) := \{k \in N : \; d_{ik}(g) < d_{ik}(g\backslash l)\}$ as the set of players whose distance to player i increases when deleting the set of links l from network g. Since distances cannot decrease when deleting links, we can rewrite marginal utility in the following way:

$$\begin{aligned} &mu_i^{CO}(g,l) \\ &= w_{ii} + \sum_{k\neq i} \delta^{d_{ik}(g)} w_{ik} - \sum_{k:ik\in g} c_{ik} - \left[w_{ii} + \sum_{k\neq i} \delta^{d_{ik}(g\setminus l)} w_{ik} - \sum_{m:im\in g\setminus l} c_{im} \right] \\ &= \sum_{k\in\kappa_i(g,l)} (\delta^{d_{ik}(g)} - \delta^{d_{ik}(g\setminus l)}) w_{ik} - \sum_{ij\in l} c_{ij}. \end{aligned}$$

Now, consider some network g, some player i and some set of player i's links $l \subseteq L_i(g)$. Suppose that $|l| \geq 2$.[18]

To show the claim, let us assume the contrary, i.e. $mu_i^{CO}(g,l) < \sum_{ij\in l} mu_i^{CO}(g,ij)$.

$$mu_i(g,l) < \sum_{ij\in l} mu_i(g,ij)$$

$$\sum_{k\in\kappa_i(g,l)} (\delta^{d_{ik}(g)} - \delta^{d_{ik}(g\setminus l)}) w_{ik} - \sum_{ij\in l} c_{ij} < \sum_{ij\in l} \left[\sum_{k\in\kappa_i(g,ij)} (\delta^{d_{ik}(g)} - \delta^{d_{ik}(g\setminus ij)}) w_{ik} - c_{ij} \right]$$

$$\sum_{k\in\kappa_i(g,l)} (\delta^{d_{ik}(g)} - \delta^{d_{ik}(g\setminus l)}) w_{ik} < \sum_{ij\in l} \sum_{k\in\kappa_i(g,ij)} (\delta^{d_{ik}(g)} - \delta^{d_{ik}(g\setminus ij)}) w_{ik} \tag{4}$$

To see that Eq. 4 cannot hold, note the following three properties of geodesic distances that were also used in Calvó-Armengol and Ilkiliç (2009):

1. $\forall k \in N$ and $\forall ij \in l$, it holds that $(\delta^{d_{ik}(g)} - \delta^{d_{ik}(g\setminus l)}) w_{ik} \geq (\delta^{d_{ik}(g)} - \delta^{d_{ik}(g\setminus ij)}) w_{ik}$.
2. For all $ij, im \in l$, it holds that $\kappa_i(g,ij) \cap \kappa_i(g,im) = \emptyset$.
3. $\bigcup_{ij\in l} \kappa_i(g,ij) \subseteq \kappa_i(g,l)$.

Thus, we conclude that $mu_i(g,l) \geq \sum_{ij\in l} mu_i(g,ij)$. □

Proposition 1 *In the symmetric connections model, g^{Tetra} is under-connected with respect to the utilitarian welfare function for any parameters δ and c, for which g^{Tetra} is pairwise stable.*

Proof We have to show that if δ and c are such that $g^{Tetra} \in PS(u_{\delta,c})$, then $\exists g' \supset g^{Tetra}$ for which $w_{\delta,c}(g') > w_{\delta,c}(g^{Tetra})$. Specifically, we show that the condition

$$c \leq \delta - \delta^8 + \delta^2 - \delta^7 + \delta^3 - \delta^6 + 2(\delta^4 - \delta^5) := ub \tag{5}$$

is necessary for stability, but sufficient for $w_{\delta,c}(g^{Tetra} \cup \{2, 13\}) > w_{\delta,c}(g^{Tetra})$. The labels of the players correspond to Fig. 1.

[18] For $|l| < 2$ the claim $mu_i(g,l) \geq \sum_{ij\in l} mu_i(g,ij)$ trivially holds.

The first part was done in Jackson and Wolinsky (1996) already. Suppose that $c > ub$, then player 1 benefits from cutting $\{1, 2\}$ (because his change in benefits is just ub).

For the second part denote by $\beta_i := \sum_{j \neq i} \delta^{d_{ij}(g^{Tetra} \cup \{2,13\})} - \sum_{j \neq i} \delta^{d_{ij}(g^{Tetra})}$ the marginal benefits for player i and by $\Delta := \sum_{i \in N} \beta_i$ the sum of marginal benefits. This allows us to write

$$w_{\delta,c}(g^{Tetra} \cup \{2, 13\}) > w_{\delta,c}(g^{Tetra}) \Longleftrightarrow \Delta > 2c. \qquad (6)$$

It is straightforward to derive that

$$\begin{aligned}
\beta_1 &= \beta_{12} = \delta^2 - \delta^4 + \delta^3 - \delta^4 \\
\beta_2 &= \beta_{13} = \delta - \delta^5 + \delta^2 - \delta^4 + \delta^2 - \delta^5 + 2(\delta^3 - \delta^4) \\
\beta_3 &= \delta^2 - \delta^5 + \delta^3 - \delta^4 + \delta^3 - \delta^5 \\
\beta_4 &= \beta_7 = \beta_9 = \beta_{15} = \delta^3 - \delta^4 \\
\beta_{14} &= \delta^2 - \delta^5 + \delta^3 - \delta^4 + \delta^3 - \delta^5,
\end{aligned}$$

and $\beta_i = 0$ for all other i.

This yields

$$\Delta = 2(\delta - \delta^5) + 4(\delta^2 - \delta^4) + 4(\delta^2 - \delta^5) + 12(\delta^3 - \delta^4) + 2(\delta^3 - \delta^5). \qquad (7)$$

To show that $\Delta > 2c$ under the condition $c \leq ub$, it is sufficient to show that $\Delta > 2ub$ holds. Recall that,

$$2ub(g) = 2(\delta - \delta^8) + 2(\delta^2 - \delta^7) + 2(\delta^3 - \delta^6) + 4(\delta^4 - \delta^5). \qquad (8)$$

Thus,

$$\Delta > 2ub \Longleftrightarrow 6\delta^2 + 12\delta^3 - 20\delta^4 - 4\delta^5 + 2\delta^6 + 2\delta^7 + 2\delta^8 > 0 \qquad (9)$$

Numerically it can be checked that (9) holds for all $\delta \in (0, 1)$ (we used Maple). □

Proof of Proposition 2 Let $c < \frac{\rho}{(\rho+2k+2)(\rho+2k)}$, then for the welfare maximizing number of links it holds that $1/2D^*(g) \geq k$. Since any network, which contains $1/2D^*(g)$ links is welfare maximizing, any network, which has less than $1/2D^*(g)$ links is under-connected. By Theorem 2 no pairwise stable network can be under-connected since u^{PR} satisfies negative externalities and concavity. Thus, any network $g \in [PST]$ has to contain at least $1/2D^*(g) \geq k$ links. □

References

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

Belleflamme P, Bloch F (2004) Market sharing agreements and collusive networks. Int Econ Rev 45(2):387–411

Bloch F, Dutta B (2009) Communication networks with endogenous link strength. Games Econ Behav 66(1):39–56
Bloch F, Jackson M (2006) Definitions of equilibrium in network formation games. Int J Game Theory 34(3):305–318
Bloch F, Jackson MO (2007) The formation of networks with transfers among players. J Econ Theory 133(1):83–110
Buechel B, Hellmann, T (2009) Under-connected and over-connected networks. Working papers 2009.38, Fondazione Eni Enrico Mattei
Buskens V, van de Rijt A (2008) Dynamics of networks if everyone strives for structural holes. Am J Sociol 114:371–407
Calvó-Armengol A, Ilkiliç R (2009) Pairwise-stability and nash equilibria in network formation. Int J Game Theory 38(1):51–79
Chakrabarti S, Gilles R (2007) Network potentials. Rev Econ Des 11(1):13–52
Dasgupta P, Stiglitz J (1980) Uncertainty, industrial structure, and the speed of r&d. Bell J Econ 11(1):1–28
Dutta B, Mutuswami S (1997) Stable networks. J Econ Theory 76(2):322–344
Goyal S, Joshi S (2006a) Bilateralism and free trade. Int Econ Rev 47(3):749–778
Goyal S, Joshi S (2006b) Unequal connections. Int J Game Theory 34(3):319–349
Goyal S, Vega-Redondo F (2007) Structural holes in social networks. J Econ Theory 137(1):460–492
Hellmann T (2009) Convexity and complementarity in network formation: implications for the structure of pairwise stable networks. Working papers 423, Bielefeld University, IMW
Hummon NP (2000) Utility and dynamic social networks. Soc Netw 22(3):221–249
Jackson MO (2003) The stability and efficiency of economic and social networks. In: Sertel M, Koray S (eds) Advances in economic design. Springer, Heidelberg
Jackson MO (2004) A survey of models of network formation: stability and efficiency. In: Demange G, Wooders M (eds) Group formation in economics: networks, clubs and coalitions. Cambridge University Press, Cambridge, UK, pp 11–49
Jackson MO, Wolinsky A (1996) A strategic model of social and economic networks. J Econ Theory 71(1):44–74
Myerson RB (1991) Game theory: analysis of conflict. Harvard University Press, Cambridge, Massachusetts
Rogers BW (2006) A strategic theory of network status. Technical report, California Institute of Technology