Networks as Public Infrastructure: Externalities, Efficiency, and Implementation*

Hans Haller†

August 30, 2014

Abstract

First a non-cooperative model of network formation is investigated where link formation is one-sided and information flow is two-way. For that model, the relationship between different notions of efficient networks is studied: Pareto optimal networks on the one hand and welfare maximizing networks on the other hand. Strategic network formation is compared with funding schemes for public goods. Second, we extend the model and review earlier findings how a pre-existing network affects existence of Nash equilibria and efficiency of Nash equilibrium outcomes of the strategic network formation game: It can foster or prohibit existence of Nash equilibria. It can improve or worsen equilibrium welfare. Finally, we treat the pre-existing network as public infrastructure and design and analyze a subscription game for the public provision of that infrastructure.

JEL Classification: C72, D85
Keywords: Network externalities; Efficiency; Strategic network formation; Network extension; Implementation

*The comments of a referee and the suggestions of an associate editor are appreciated.
†Department of Economics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA. *Email: haller@vt.edu.*

1 Introduction

The standard competitive equilibrium model — of pure exchange or with production — assumes private goods without externalities. Allowing for externalities renders the analysis more cumbersome and, as a rule, affects the welfare conclusions. Introduction of public goods leads to the Samuelsonian characterization of Pareto optimal allocations and implementation of the latter by means of mechanisms different from the standard market mechanism, for instance by means of personalized or Lindahl prices. Both the treatment of externalities and of public goods are considered special fields of microeconomic theory. In contrast, externalities are the essence of strategic game theory and its applications, like those in industrial organization and other areas of economics. A strategic game where none of the players is affected by the actions of others is justifiably termed "inessential".

Following the seminal contributions of Jackson and Wolinsky (1996) and Bala and Goyal (2000) and the less well known earlier example by Myerson (1991, page 448), a sizeable game-theoretic literature on network formation has emerged. The aim of the current paper is to address several issues that arise when (a) a network is viewed as a public infrastructure and (b) the network or part of it is the outcome of a game in strategic form. The main issues to be addressed or at least touched upon in a first pass are: network externalities (and public good features of networks), efficiency, and implementation; and those are obviously related.

In accordance with my prior work and for ease of exposition, I shall adhere to the purely noncooperative approach of Bala and Goyal (2000) where addition and deletion of links are unilateral decisions of the player from whom the respective links originate.[1] More specifically, I shall focus on the so-called two-way flow connections model à la Galeotti, Goyal, and Kamphorst (2006) both as a benchmark and an illustrative prototype. That model incorporates cost and value heterogeneity. A player receives valuable information from others not only through direct links, but also via indirect links. The player incurs the cost of the direct links she initiates. Before continuing the detailed description of the model, I make a digression on each of the three themes, network externalities (and public good features of networks), efficiency and implementation.

[1] In contrast, pairwise stability à la Jackson and Wolinsky (1996) treats addition of a link in a network as a bilateral decision by the two players involved, whereas severance of a link constitutes a unilateral decision.

Network Externalities. A strategic network formation game like any interesting strategic game inherently involves externalities. So-called "network externalities" prove relevant for network formation as well as network utilization, for both positive and normative analysis. A classic and popular example for positive network externalities or "network effects" is the telephone. Having a telephone is only of interest to the user if there is a significant number of other users one can reach (and wants to reach) through the same network. If additional users join the network, then they increase the value of the telephone to the already existing users, even if that is not their intention. The two-way flow connections model shares this property: If a player creates a direct link to another player, the initiating player receives valuable information from the other player. In addition, the first player receives valuable information from third players via indirect links and other players receive valuable information from him through direct and indirect links. A player can only benefit and never gets harmed when others create links.

The telephone also illustrates the possibility of negative network externalities or "congestion effects". If too many users attempt to make calls simultaneously, service may be disrupted, delayed or noisy. This was particularly the case before the digital age because of overloaded lines and relays, but can still occur nowadays. Modifications of the two-way flow connections model can accommodate negative network externalities. I will briefly comment on such model variations in Section 7. However, the emphasis of the paper is going to be on positive network externalities.

Local Public Good Aspects of Networks. If in the standard two-way flow connections model (without decay and with perfectly reliable links), an isolated player forms a link to another player, then the first player gets access to all the information available within the connected component of the second player. He would receive the same information when establishing instead a link to another player of that connected component. However, the costs of link formation may be different. Investing in one of these links is like paying a fee to enter the particular component where the entry fee may differ depending on which link is used. Whether it is worth paying the fee at all may depend on the size and composition of the component. This resembles the decision of joining a club. Only when the club has a suitable membership, in terms of numbers or composition, or offers sufficiently attractive club goods, is it worth to pay the membership fee. In Section 4, we delineate the differences and parallels between strategic network formation and funding schemes for public goods.

Efficiency. A network is efficient (in the narrow traditional sense) if the aggregate payoff resulting from the network is maximal among all networks. Traditionally, efficiency plays a prominent role as a major performance criterion for network designers or planners, especially in cost-benefit analysis when payoffs are monetary. It constitutes an important benchmark for network performance even when network formation is decentralized and structured as a strategic game. In economics, the term "efficiency" is often used in a broader sense, as a synonym for Pareto optimality. An efficient network in the traditional sense is necessarily Pareto optimal. That is, there is no other network that provides a higher payoff to some player(s) and no lower payoff to any player. But a Pareto optimal network need not be efficient in the traditional sense. There exist even model specifications with Pareto optimal networks that do not maximize any utilitarian social welfare function. Section 3 provides a more systematic investigation of the relationship between notions of efficient networks: Pareto optimal networks on the one hand and welfare maximizing networks on the other hand.

Given that it matters to a player how many and which players are already in the network (in the component) when the player decides whether to form a link to an existing component, two extreme outcomes are possible. In case the players in the component do not constitute a critical mass, then not only will the player decide against joining, but the component may not come about in the first place. The empty network is a possible equilibrium outcome of the game. Otherwise, if a component of sufficient magnitude materializes, then everybody wants to be connected to that component and a minimally connected network may be the equilibrium outcome of the game. Frequently, minimally connected networks turn out to be efficient whereas the empty network often is not. In contrast, the much investigated funding scheme for public goods, voluntary contributions, tends to produce very inefficient outcomes both in theory and in experiments. In Section 4, we compare and reconcile the different welfare implications in strategic network formation and funding of public goods.

Implementation. In Haller (2012), I drew a distinction between a preexisting, possibly publicly provided network $\mathfrak{g}$ and its extension via a strategic network formation game where links are added to the existing network. There I investigated how the pre-existing network effects the final outcome, mainly comparing situations without and with some prior network. I found among other things: First, a non-trivial $\mathfrak{g}$ can prove efficiency enhancing, leading to an efficient outcome while an inefficient network may result when $\mathfrak{g}$ is the empty network with no links. Second, to the opposite, a non-trivial $\mathfrak{g}$ can prove efficiency reducing. Further, a non-trivial $\mathfrak{g}$ can be stabilizing, yielding existence of a Nash network (i.e., a network that is a Nash equilibrium outcome of the strategic network formation game) in situations where Nash networks do not exist when $\mathfrak{g}$ is the empty network. Finally, under different circumstances, a non-trivial $\mathfrak{g}$ can prove destabilizing. Since the focus of the present paper lies on efficiency, the first two effects are of primary interest here: the efficiency enhancing or efficiency reducing effect of a publicly provided network $\mathfrak{g}$. Those two possibilities suggest the question which network $\mathfrak{g}$ would be provided if the construction of $\mathfrak{g}$ was treated as a public infrastructure project. In Section 6, I model and analyze the choice between a newly provided public infrastructure (network), followed by a network extension game, and no provision of a public infrastructure (network), followed by a standard network formation game. The first-step binary choice problem is modeled as a "subscription game" for a public project.

Model Details. A network obtains as the outcome of a game in strategic form. The basic model is the two-way flow connections model à la Galeotti, Goyal, and Kamphorst (2006). The players coincide with the nodes of the network and unilaterally form links. A player receives valuable information from others not only through direct links, but also via indirect links. The player incurs the cost of the direct links she initiates. The solution concept is Nash equilibrium. In general, the term Nash network refers to a network arising as the Nash equilibrium outcome of a network formation game. I shall use **Nash network** as a synonym for Nash equilibrium, since as in the literature in the tradition of Bala and Goyal (2000), a strategy profile can be identified with the resulting network.

In the modified model of Section 5, there is a pre-existing network or infrastructure $\mathfrak{g}$ which is extended as the result of a strategic network formation game. In Section 6, the infrastructure $\mathfrak{g}$ is treated as a public project and implemented as the outcome of a "subscription game". That constitutes the first stage of a two-stage game, with the strategic network formation game as the second stage.

Main Results. Examples 1 and 2 and Proposition 1 deal with the different concepts of efficient networks. Specifically, Example 2 demonstrates the possibility of a Pareto optimal network that does not maximize any utilitarian social welfare function. Example 3 illustrates that for some parameter constellations, the unique Nash network is efficient and for other parameters, both an efficient and an inefficient Nash network can exist. In contrast, the standard "voluntary contribution game" for public goods tends to yield inefficient outcomes. We identify the source of this difference and argue that welfare conclusions similar to those in Example 3 can be achieved if the public good becomes a local public good or club good and a "fixed fee game" instead of the "voluntary contribution game" is played.

Next some of the key findings of Haller (2012) on network extension are reviewed: existence of Nash networks, a network analogue of the second welfare theorem, the efficiency enhancing or efficiency reducing effect of a publicly provided network infrastructure. In Section 6, implementation of that infrastructure via a "subscription game" is considered, with a binary choice between a particular costly infrastructure and none (the empty network or infrastructure). As a rule, of the two infrastructures the one with the more efficient two-stage outcome is going to be implemented. However, there is one noteworthy exception: There can be public provision of a costly network infrastructure even though aggregate welfare is diminished after the creation of that infrastructure. The beneficiaries may be willing to shoulder the entire cost of its construction while others are significantly harmed and opposed to it.

Outline. Section 2 develops the basic model of strategic network formation. Section 3 elaborates on the concept of efficiency of networks. Section 4 delineates the differences and parallels between the strategic network formation game and a voluntary contribution game for public goods. In Section 5, the model is modified to allow for a pre-existing or publicly provided network (infrastructure). Section 6 addresses the implementation of a public network (infrastructure). Section 7 offers some final remarks.

2 Basic Model

The basic model is identical with the benchmark model of Haller (2012) and the two-way flow connections model à la Galeotti, Goyal, and Kamphorst (2006) that incorporates cost and value heterogeneity. I adopt the notation of Haller, Kamphorst and Sarangi (2007) for the case of perfectly reliable links. The model serves several purposes. First, it constitutes the foundation of the more general models in Sections 5 and 6. Second, it facilitates the discussion of efficiency in Sections 3 and 4. Third, it serves as a benchmark and allows to assess the effects of a pre-existing and possibly publicly provided network or infrastructure.

Let $n \geq 3$. $N = \{1, \ldots, n\}$ denotes the set of players with generic elements i, j, k. N also constitutes the set of nodes of the network to be formed. For ordered pairs $(i, j) \in N \times N$, the shorthand notation ij is used and for non-ordered pairs $\{i, j\} \subset N$ the shorthand $[ij]$ is used. The symbol $\subset$ for set inclusion permits equality. The model is specified by two families of parameters, indexed by ij, with $i \neq j$:

- Cost parameters $c_{ij} > 0$.
- Value parameters $V_{ij} > 0$.

In case $c_{ij} \neq c_{kl}$ ($V_{ij} \neq V_{kl}$) for some $ij \neq kl$, the model exhibits **cost (value) heterogeneity**; otherwise, it exhibits **cost (value) homogeneity**. Following Derks and Tennekes (2009), we say that costs are **owner-homogeneous** if for each player i, there exists $c_i > 0$ such that $c_{ij} = c_i$ for all $j \neq i$. This condition is also considered in Galeotti (2006), Galeotti *et al.* (2006), Billand *et al.* (2008), Derks and Tennekes (2009), and Haller (2012).

We only consider pure strategies. A pure strategy for player i is a vector $g_i = (g_{i1}, \ldots, g_{ii-1}, g_{ii+1}, \ldots, g_{in}) \in \{0, 1\}^{N \setminus \{i\}}$. The set of all pure strategies of agent i is denoted by $\mathcal{G}_i$. It consists of 2^{n-1} elements. The joint strategy space is given by $\mathcal{G} = \mathcal{G}_1 \times \cdots \times \mathcal{G}_n$.

There is a one-to-one correspondence between the set of joint strategies $\mathcal{G}$ and the set of all directed graphs or networks with vertex set N. Namely, to a strategy profile $g = (g_1, \ldots, g_n) \in \mathcal{G}$ corresponds the graph $(N, E(g))$ with edge or link set $E(g) = \{(i, j) \in N \times N |\, i \neq j,\, g_{ij} = 1\}$. In the sequel, we shall identify a joint strategy g and the corresponding graph and use the terminology **directed graph** or **directed network** g. Since our aim is to model network formation, $g_{ij} = 1$ is interpreted to mean that a direct link between i and j is initiated by player i (edge ij is formed by i) whereas $g_{ij} = 0$ means that i does not initiate the link (ij is not formed). Regardless of what player i does, player j can set $g_{ji} = 1$, i.e., initiate a link with i, or set $g_{ji} = 0$, i.e., not initiate a link with i.

Benefits. A link between agents i and j potentially allows for **two-way (symmetric) flow of information**. Accordingly, the benefits from network g are derived from its closure $\overline{g} \in \mathcal{G}$, defined by $\overline{g}_{ij} := \max\{g_{ij}, g_{ji}\}$ for $i \neq j$. Moreover, a player receives information from others not only through direct links, but also via indirect links. To be precise, information flows from player j to player i, if i and j are linked by means of a path in $\overline{g}$ from i to j. A **path** of length m in $f \in \mathcal{G}$ from player i to player $j \neq i$, is a finite sequence $i_0, i_1, \ldots, i_m$ of pairwise distinct players such that $i_0 = i$, $i_m = j$, and $f_{i_k i_{k+1}} = 1$ for $k = 0, \ldots, m-1$. Let us denote

$$N_i(f) = \{j \in N |\, j \neq i, \text{ there exists a path in } f \text{ from } i \text{ to } j\},$$

the set of other players whom player i can access or "observe" in the network f. Information received from player j is worth V_{ij} to player i. Therefore, player i's benefit from a network g with perfectly reliable links and two-way flow of information is (as in Galeotti, Goyal, and Kamphorst (2006)):

$$B_i(g) = B_i(\overline{g}) = \sum_{j \,\in\, N_i(\overline{g})} V_{ij}.$$

Notice that $\overline{g}$ belongs to the set $\mathcal{H} = \{h \in \mathcal{G} |\; h_{ij} = h_{ji} \text{ for } i \neq j\}$. In turn, there is a one-to-one correspondence between the elements of $\mathcal{H}$ and the non-directed networks (graphs) with node set N. Namely, for $h \in \mathcal{H}$ and $i \neq j$, $[ij]$ is an edge of the corresponding non-directed network if and only if $h_{ij} = h_{ji} = 1$. In the sequel, we shall identify h with the corresponding non-directed network. In that case, the notation $[ij] \in h$ stands for "$[ij]$ is an edge of h", that is h is given by its set of edges. Accordingly, for $k \in \mathcal{H}$, $k \subset h$ means that k is a subnetwork of h.

Costs. Player i incurs the cost c_{ij} when she initiates the direct link ij, i.e., if $g_{ij} = 1$. Hence i incurs the total costs

$$C_i(g) = \sum_{j \neq i} g_{ij} c_{ij}$$

when the network g is formed.

Payoffs. Player i's payoff from the strategy profile g is the net benefit

$$\Pi_i(g) = B_i(g) - C_i(g). \tag{1}$$

Nash Networks. Given a network $g \in \mathcal{G}$, let g_{-i} denote the network that remains when all of agent i's links have been removed so that $g_{-i} \in \mathcal{G}_{-i} \equiv \prod_{j \neq i} \mathcal{G}_j$. Clearly $g = g_i \oplus g_{-i}$ where the symbol $\oplus$ indicates that g is formed by the union of links in g_i and g_{-i}. A strategy g_i is a **best response** of agent i to g_{-i} if

$$\Pi_i(g_i \oplus g_{-i}) \geq \Pi_i(g_i' \oplus g_{-i}) \text{ for all } g_i' \in \mathcal{G}_i.$$

Let $BR_i(g_{-i})$ denote the set of agent i's best responses to g_{-i}. A network $g = (g_1, \ldots, g_n)$ is said to be a **Nash network** if $g_i \in BR_i(g_{-i})$ for each i, that is if g is a Nash equilibrium of the strategic game with normal form $(N, (\mathcal{G}_i)_{i \in N}, (\Pi_i)_{i \in N})$. A strict Nash network is one where agents are playing strict best responses.

Some Graph-theoretic Concepts. We now introduce some definitions of a more graph-theoretic nature. The network with no links is called the **empty network** and will be denoted e. A network g is said to be **connected** if there is a path in $\overline{g}$ between any two agents i and j. A connected network g is **minimally connected**, if it is no longer connected after the deletion of *any* link.

A set $C \subset N$ is called a **component** of g if there exists a path in $\overline{g}$ between any two different agents i and j in C and there is no strict superset C' of C for which this holds true. For each network g, the components of g form a partition of the player set (node set, vertex set) N into non-empty subsets. Each isolated point $i \in N$ in g, that is a player or node i with $g_{ij} = g_{ji} = 0$ for all $j \neq i$, gives rise to a singleton component $\{i\}$. In particular, the components of the empty network are the sets $\{i\}, i \in N$. N is the only component of g if and only if g is connected. If C is a component of the network g, we denote by g^C the network induced by g on the set of nodes C, that is $g_{ij}^C = g_{ij}$ for $i, j \in C$, $i \neq j$. A network g is **minimal**, if g^C is minimally connected for every component C of g. Both Nash networks and efficient networks (to be defined next) are minimal. Minimally connected networks are both connected and minimal.

3 Efficient Networks

Let $W_0 : \mathcal{G} \to \mathbb{R}$ be defined as $W_0(g) = \sum_{i=1}^{n} \Pi_i(g)$. A network $\widehat{g}$ is **efficient (in the narrow traditional sense)** if $W_0(\widehat{g}) \geq W_0(g)$ for all $g \in \mathcal{G}$. Efficiency is a major performance criterion for network designers or planners and plays a prominent role in the traditional network literature. It is most attractive for cost-benefit analysis when payoffs are monetary and side-payments between players are feasible. Efficiency constitutes an important benchmark for network performance even when network formation is decentralized and structured as a strategic game. In economics, the term "efficiency" is often used in a broader sense, as a synonym for Pareto optimality. An efficient network in the traditional sense is necessarily Pareto optimal. That is, there is no other network that provides a higher payoff to some player(s) and no lower payoff to any player. But not every Pareto optimal network has to be efficient in the traditional sense as the following example demonstrates.[2]

Example 1 (Efficiency and Social Welfare)
Let $N = \{1, 2, 3\}$. Further set $V_{ij} = 1$ for all $i \neq j$; $c_{2j} = 1.6$ and $c_{j2} = 1.8$ for $j \neq 2$; and $c_{ij} = 3$ otherwise. Then the empty network and $g' = \{12, 32\}$, the periphery-sponsored star with center 2, are the two Nash networks. Neither one is efficient in the traditional sense. The center-sponsored star with center 2, $\widehat{g} = \{21, 23\}$ is the only efficient network in the traditional sense. However, g' is Pareto optimal and maximizes the utilitarian social welfare function $W(g) = \Pi_1(g) + 2\Pi_2(g) + \Pi_3(g)$. □

In the example, the Pareto optimal network g' does not maximize the social welfare function W_0 with equal weights, but it maximizes another utilitarian social welfare function on $\mathcal{G}$. This raises the question whether in general, for every Pareto optimal network $g' \in \mathcal{G}$, there exists a social welfare function $W : \mathcal{G} \to \mathbb{R}$ of the form $W(g) = \sum_{i=1}^{n} a_i \Pi_i(g)$ so that

[2] The discrepancy between Pareto optimality and efficiency has also been pointed out by Haller and Sarangi (2005).

$(a_1, \dots, a_n) \gg 0$ and g' maximizes W on $\mathcal{G}$.

Pareto optimal networks are minimal, but not necessarily minimally connected. We obtain some partial answers in case all Pareto optimal networks are minimally connected:

Proposition 1 *Suppose all Pareto optimal networks are minimally connected.*

(a) *For generic utilitarian welfare weights, the corresponding utilitarian social welfare function has a unique maximizer.*

(b) *If costs are not owner-homogeneous, there can exist Pareto optimal networks that do not maximize any utilitarian social welfare function.*

(c) *In case of owner-homogeneous costs, there exists a utilitarian social welfare function W^* such that all minimally connected networks maximize W^* and are Pareto optimal.*

A social planner may find it desirable that everybody has access to the network, but that there are no redundant links. This would favor minimally connected networks, even if they are not Pareto optimal. In case they are, the choice of a particular welfare weight vector $(a_1, \dots, a_n) \gg 0$ and the corresponding minimal spanning tree(s) may or may not select a specific Pareto optimum: Assertion (a) says that generically, a unique Pareto optimal network will be picked. Assertion (b) says that in some instances, none of the Pareto optima will be picked, regardless of the choice of positive weight vector, the analogue of a severe non-convexity in finite economies. Assertion (c) says that sometimes, a particular weight vector does not discriminate between Pareto optima, the analogue of a flat Pareto frontier (in utility space) in finite economies. That possibility should not be ignored because the question which welfare weights or which Pareto optima would or should be chosen by a policy maker (or represent a social consensus), does not have an obvious answer.

PROOF. If g is minimally connected and $W(g) = \sum_{i=1}^{n} a_i \Pi_i(g)$ is a utilitarian social welfare function, then

$$W(g) = \sum_{i=1}^{n} a_i \sum_{j \neq i} V_{ij} - \sum_{i=1}^{n} a_i \sum_{j \neq i} g_{ij} c_{ij}. \tag{2}$$

Hence maximization of $W(g)$ on the set of minimally connected networks amounts to minimization of $\sum_{i=1}^n a_i \sum_{j \neq i} g_{ij} c_{ij}$ on $\mathcal{G}$ subject to the constraint that g is minimally connected. In the literature on combinatorial optimization, computer science, and operations research, the latter is known as the **minimum spanning tree problem** with weights $w_{ij} = a_i c_{ij}$.

(a) $a_i c_{ij} \neq a_k c_{kl}$ for $ij \neq kl$ unless $a_i = a_k \cdot c_{kl}/c_{ij}$. Hence with the exception of $(a_1, \ldots, a_n) \gg 0$ belonging to a finite number of hyperplanes in $\mathbb{R}^n$, the weights $w_{ij} = a_i c_{ij}$ satisfy $w_{ij} \neq w_{kl}$ for $ij \neq kl$. But if the weights w_{ij} are distinct across pairs ij, the minimum spanning tree is unique. See, for instance, Property 2 and its proof in Gallager *et al.* (1983).

(b) See Example 2 below.

(c) In case of owner-homogeneous costs, for each player i, there exists $c_i > 0$ such that $c_{ij} = c_i$ for all $j \neq i$. Set $a_i^* = \prod_{j \neq i} c_j$ for each $i \in N$ and $b = \prod_i c_i$. Take the utilitarian social welfare function $W^*(g) = \sum_{i=1}^n a_i^* \Pi_i(g)$. Then $a_i^* c_i = b$ for all i and $W^*(g) = \sum_{i=1}^n a_i^* \sum_{j \neq i} V_{ij} - (n-1)b$ for all minimally connected networks g. If there exists a network g with $W^*(g) > \sum_{i=1}^n a_i^* \sum_{j \neq i} V_{ij} - (n-1)b$, then there exists a maximizer of W^* that is not minimally connected but Pareto optimal, contradicting the assumption that all Pareto optimal networks are minimally connected. Hence all minimally connected networks are maximizers of W^* and Pareto optimal. ■

Example 2 (Pareto Optimality and Social Welfare)
In this example, all Pareto optimal networks are minimally connected but some do not maximize any utilitarian social welfare function.

Let $N = \{1, 2, 3, 4, 5\}$. Further set $V_{ij} = 8$ for all $i \neq j$; $c_{41} = c_{51} = 2, c_{42} = c_{52} = 4, c_{43} = 6, c_{53} = 5$; and $c_{ij} = 30$ otherwise. In this example, there exist at least nine Pareto optimal networks and all Pareto optimal networks are minimally connected. First, let us show that every Pareto optimal network is minimally connected. As an immediate consequence of the definitions, a Pareto optimal network is minimal. It remains to be shown that it is connected. If g is a network and $i \in \{4, 5\}$ and $j \in \{1, 2, 3\}$ belong to different components of g, then $\Pi_i(g \oplus \{ij\}) > \Pi_i(g)$ and $\Pi_{i'}(g \oplus \{ij\}) \geq \Pi_{i'}(g)$ for all $i' \in N$. Hence in this example, a Pareto optimal network can have only one component, that is, it is connected.

Let us focus next on four specific Pareto optimal networks. We label these networks g^b, g^d, g^p, g^q such that $x = (\Pi_4(g^x), \Pi_5(g^x))$ for $x = b, d, p, q$. In each of the four networks, both 4 and 5 initiate a link to 1. Moreover, for $j = 2, 3$ either 4 or 5 initiates a link to j. In all four networks, players $j = 1, 2, 3$

don't initiate any links and receive payoffs $\Pi_j(g^x) = 32$. The payoff pairs $x = (\Pi_4(g^x), \Pi_5(g^x))$ are $b = (20, 30), d = (24, 26), p = (26, 25), q = (30, 21)$. The five periphery sponsored stars constitute additional Pareto optimal networks.

Now consider an arbitrary utilitarian social welfare function W with weights $(a_1, \ldots, a_n) \gg 0$.

- If $a_4 < a_5$, then
  $a_4 \cdot b_4 + a_5 \cdot b_5 = 20 \cdot a_4 + 30 \cdot a_5 > 24 \cdot a_4 + 26 \cdot a_5 = a_4 \cdot d_4 + a_5 \cdot d_5$;
  hence, $W(g^b) > W(g^d)$ and g^d does not maximize W.

- If $a_4 \geq a_5$, then
  $a_4 \cdot p_4 + a_5 \cdot p_5 = 26 \cdot a_4 + 25 \cdot a_5 > 24 \cdot a_4 + 26 \cdot a_5 = a_4 \cdot d_4 + a_5 \cdot d_5$;
  hence, $W(g^p) > W(g^d)$ and g^d does not maximize W.

It follows that g^d does not maximize W in any case. Thus, g^d proves to be a Pareto optimal network that does not maximize any utilitarian social welfare function. □

The simplest type of the model with homogeneous costs and values proves very instructive for welfare comparisons.

Example 3 (Homogeneous Costs and Values)
Let $c_{ij} = c > 0$ and $V_{ij} = V = 1$ for all $i \neq j$.

(a) $c < 1$: Then the Nash networks and efficient networks coincide and consist of the minimally connected networks.

(b) $n - 2 < c < n - 1$: The empty network and the periphery-sponsored stars are the only Nash networks. The former is inefficient whereas the latter are efficient (as are all minimally connected networks).

The two cases (a) and (b) are not exhaustive, but the most interesting ones. □

4 Comparison with Voluntary Contribution Game

Voluntary contributions to a public good are frequently modeled as follows. There is one public good, provided in quantities $y \geq 0$ and one private good consumed in quantities c. The unit of measurement of each good is chosen such that the unit price is 1. There are $n > 1$ consumers in the community or society, indexed $i = 1, \dots, n$, each with an endowment ω_i of the private good, and with utility function $u(c_i, y) = c_i + y$ where y is the amount of the public good and c_i is i's consumption of the private good.

Voluntary Contribution Game. The public good is produced by means of a technology described by a strictly concave and differentiable production function $f : \mathbb{R}_+ \to \mathbb{R}_+$, satisfying $f(0) = 0$ and $f' > 0$. The private good also serves as input in the production of the public good. If the amount $x \geq 0$ is used as input, $f(x)$ units of the public good will be produced. Each consumer i voluntarily contributes an amount $x_i \geq 0$ of the private good to the production of the public good. With total contribution $x = \sum_{i=1}^{n} x_i$, the outcome is $y = f(x) = f(\sum_i x_i)$, with resulting utilities $u(\omega_i - x_i, y) = f(\sum_j x_j) + \omega_i - x_i$. Neglecting the constant terms ω_i, one obtains a strategic game in normal form $(I, (X_i)_{i\in I}, (U_i)_{i\in I})$, the **"voluntary contribution game"**, with player set $I = \{1, \dots, n\}$, strategy set $X_i = \mathbb{R}_+$ and payoff function $U_i(x_1, \dots, x_n) = f(\sum_j x_j) - x_i$ for $i \in I$.

I shall disregard conceivable variations and complications of the model: non-negativity constraints for c_i and ω_i; more convoluted utility function(s); heterogeneity in consumer preferences; money instead of a consumption good; a cost function in lieu of a production function; multiple public goods.

Equilibrium Analysis. The Nash equilibrium analysis of the game $(I, (X_i)_{i\in I}, (U_i)_{i\in I})$ is straightforward. If $f'(0) \leq 1$, then $x_1^* = x_2^* = \dots = x_n^* = 0$ is the unique Nash equilibrium (in pure strategies). If $f'(0) > 1$, then there exists a unique $x^* > 0$ with $f'(x^*) = 1$ and the set of Nash equilibria consists of all $(x_1^*, \dots, x_n^*) \in X_1 \times \dots \times X_n$ such that $\sum_i x_i^* = x^*$. For later reference, denote $y^* = f(x^*)$, the amount of public good that obtains in a Nash equilibrium.

Welfare Analysis. The standard welfare analysis of the voluntary contribution game proceeds as follows. Consider the aggregate social welfare function $W : X_1 \times \dots \times X_n \to \mathbb{R}$ given by $W = \sum_i U_i$. Then for $(x_1, \dots, x_n) \in X_1 \times \dots \times X_n$, $W(x_1, \dots, x_n) = nf(\sum_j x_j) - \sum_j x_j$. Hence W is maximized at $(x_1, \dots, x_n)$ if and only if $\sum_j x_j = \widehat{x}$ where $\widehat{x}$ is the unique solution of $f'(x) = 1/n$ (assuming $f'(0) > 1/n$). Such $(x_1, \dots, x_n)$ and the corresponding $\widehat{y} = f(\widehat{x})$ are referred to as efficient outcomes or more specifically, efficient contribution profiles and the efficient amount of the public good, respectively. Clearly, $\widehat{x} > x^*$ and $\widehat{y} > y^*$. Moreover, if $(x_1^*, \ldots, x_n^*)$ is a Nash equilibrium profile, then there exists an efficient contribution profile $(x_1, \ldots, x_n)$ such that for each $i \in I$, $U_i(x_1, \ldots, x_n) > U_i(x_1^*, \ldots, x_n^*)$. Namely, for each $i \in I$, choose $x_i' \in X_i$ such that $\widehat{y} - x_i' = U_i(x_1^*, \ldots, x_n^*) = y^* - x_i^*$. Then $x_i' > x_i^*$, since $\widehat{y} > y^*$. Let $x' = \sum_j x_j'$. Then $n\widehat{y} - x' = ny^* - x^* < n\widehat{y} - \widehat{x}$. Hence $x' > \widehat{x} > x^*$ and there exists $\lambda \in (0,1)$ such that $\lambda x' + (1-\lambda)x^* = \widehat{x}$. Set $x_i = \lambda x_i' + (1-\lambda)x_i^*$ for $i \in I$. Then $(x_1, \ldots, x_n)$ is an efficient contribution profile and $U_i(x_1, \ldots, x_n) > U_i(x_1^*, \ldots, x_n^*)$ for all $i \in I$. That is, every Nash equilibrium profile is strictly Pareto dominated by some efficient contribution profile. In particular, the symmetric Nash equilibrium profile $(x^*/n, \ldots, x^*/n)$ is strictly Pareto dominated by the symmetric efficient contribution profile $(\widehat{x}/n, \ldots, \widehat{x}/n)$. However, not every Nash equilibrium profile may be dominated by the symmetric efficient contribution profile as will be demonstrated in Example 5.

The main differences between the strategic network formation game of Example 3 and the voluntary contribution game are two-fold: First, an isolated node does not benefit from a network formed by other nodes. Only by incurring a link cost does the player get access to the network and enjoy the benefits from direct links as well as the positive externalities caused by indirect links. In contrast, the pure public good is non-exclusive. A consumer reaps the benefit of the public good regardless of her contribution. Second, link costs are discrete whereas the voluntary contribution is a continuous choice. This suggests that efficient provision of the public good can be achieved if it becomes a local public good and excludable like a club good. That means that a consumer can be denied access to the public good unless she pays the fixed fee, say $\widehat{x}/n$ to help finance the efficient amount of the public good.

Fixed Fee Game. Let us examine a corresponding strategic game in normal form $(I, (S_i)_{i \in I}, (\Pi_i)_{i \in I})$, the **"fixed fee game"**: The player set is again $I = \{1, \ldots, \}$. Player i has the binary strategy set $S_i = \{0, \widehat{x}/n\}$. For a strategy profile $s = (s_1, \ldots, s_n) \in S = S_1 \times \ldots \times S_n$, let $k(s) = |\{i \in I : s_i = \widehat{x}/n\}|$, the number of players who opt for the public good (who "join the club"). Then the payoff of player i is $\Pi_i(s) = f(k(s) \cdot \widehat{x}/n) - \widehat{x}/n$ if $s_i = \widehat{x}/n$ and $\Pi_i(s) = 0$ if $s_i = 0$. If $f'(0) > 1/n$, then (a) $f(\widehat{x}) - \widehat{x}/n > 0$; (b) $(s_1^*, \ldots, s_n^*) = (\widehat{x}/n, \ldots, \widehat{x}/n)$ is a Nash equilibrium and the only Nash equilibrium where a player opts for the public good. This means that if at an equilibrium, some players opt for the public good, then all players opt for the public good. Typically, there is a second Nash equilibrium, $s^{**} = (0, \ldots, 0)$ where none of the players opt for the public good. Thus, the possible equilibrium outcomes resemble those of the strategic network formation game of Example 3.

Example 4. Let $n > 5$ and $f(x) = 2\sqrt{x}$ for $x \geq 0$. Then $x^* = 1$ and $\widehat{x} = n^2$. $s^* = (s_1^*, \ldots, s_n^*) = (n, \ldots, n)$ and $s^{**} = (0, \ldots, 0)$ are the only Nash equilibria in the fixed fee game. If more than $n/4$ players opt for the public good, then it is a best response of every player to opt for the public good. If less than $n/4$ players opt for the public good, then it is a best response of any player to forego the public good. This is the analogue of case (b) in Example 3. □

Example 5. Let $n = 3$ and $f(x) = 2x$ for $x \leq 1$, $f(x) = 4 - 2/x$ for $x \geq 1$. Then $f'(x) = 2$ for $x \leq 1$, $f'(x) = 2/x^2$ for $x \geq 1$. This yields $x^* = \sqrt{2}$ and $\widehat{x} = \sqrt{2n}$.

Next consider the Nash equilibrium profile $(x_1^*, x_2^*, x_3^*) = (0, 0, \sqrt{2})$. Then $U_1(x_1^*, x_2^*, x_3^*) = 4 - \sqrt{2}$ whereas $U_1(\widehat{x}/n, \ldots, \widehat{x}/n) = 4 - \frac{2}{\sqrt{2n}} - \sqrt{\frac{2}{n}} = 4 - 2\sqrt{\frac{2}{n}}$. Now $n = 3$ implies $2 > \sqrt{n}$ and $2\sqrt{\frac{1}{n}} > 1$. Consequently, $2\sqrt{\frac{2}{n}} > \sqrt{2}$ and $U_1(x_1^*, x_2^*, x_3^*) > U_1(\widehat{x}/n, \ldots, \widehat{x}/n)$. This shows that there exist Nash equilibrium profiles which are not dominated by the symmetric efficient contribution profile.

Finally, $\widehat{x}/n = \sqrt{2/n} = \sqrt{2/3} < 1$ implies $f(\widehat{x}/n) = 2\widehat{x}/n > \widehat{x}/n$. It follows that $(s_1^*, \ldots, s_n^*) = (\widehat{x}/n, \ldots, \widehat{x}/n) = (\sqrt{2/3}, \ldots, \sqrt{2/3})$ is a Nash equilibrium in strictly dominant strategies in the fixed fee game. This is the analogue of case (a) in Example 3. □

5 Network Extension

A pre-existing network or infrastructure $\mathfrak{g} \in \mathcal{G}$, can be incorporated into the model in two ways: 1. via restrictions on players' strategy choices or 2. via a modification of the players' payoff functions. We opt for the second way, which proves more tractable in our context. For $g, h \in \mathcal{G}$, $h \oplus g \in \mathcal{G}$ denotes the network whose set of links is the union of the links in h and the links in

g. Given $\mathfrak{g}$, we define for each $i \in N$ a payoff function $\Pi_i(\mathfrak{g}; \cdot) : \mathcal{G} \to \mathbb{R}$ by

$$\Pi_i(\mathfrak{g}; g) = B_i(\mathfrak{g} \oplus g) - C_i(g) \text{ for } g \in \mathcal{G}. \tag{3}$$

This formulation treats $\mathfrak{g}$ free of costs which is of course implausible. Under the separable cost assumption of the connections model, the cost of providing $\mathfrak{g}$ can be neglected as long as merely the formation of a network g given $\mathfrak{g}$ is considered. We account for the cost of $\mathfrak{g}$ when performing comparative statics and welfare analysis.

In the extended model, one ought to distinguish between a Nash equilibrium g^* given the infrastructure $\mathfrak{g}$ and the outcome $\mathfrak{g} \oplus g^*$, the entire available network. Notice that $\mathfrak{g}$ and g^* are disjoint because of $c_{ij} > 0$ for all ij.

Next I summarize some of the earlier findings. Proposition 2 is relevant in the context of Section 6.

Proposition 2 (Haller (2012), Proposition 1) *Consider a strategic model of network formation with payoff functions* $\Pi_i(\mathfrak{g}; g)$, $g \in \mathcal{G}$, $i \in N$. *Suppose that costs are owner-homogeneous. Then there exists a Nash network* g^*.

If a Pareto optimal network $\mathfrak{g}$ is already in place and Pareto optimal outcomes are desired, then there is no point to add further links to $\mathfrak{g}$. Indeed, Proposition 2 of Haller (2012) states that in equilibrium, no additional links — which might undo Pareto optimality — are formed. This result is reminiscent of the second welfare theorem for standard pure exchange economies. Notice, however, that a publicly provided or pre-existing infrastructure $\mathfrak{g}$ is not necessarily Pareto optimal. If so, the question is what are its effect on equilibrium welfare. The effect of a non-trivial pre-existing or publicly provided network or infrastructure can be beneficial (Example 7) or detrimental (Example 6).[3] Example 6 demonstrates that a publicly provided infrastructure can not only be too costly, but can also induce players to choose costlier links. It further illustrates the intriguing case (bc) of Proposition 3. Although the outcome proves inefficient, in case (bc) some individuals benefit so much from the publicly provided infrastructure that they are willing to fund its construction on their own, to the detriment of others.

[3]There can also be a stabilizing or destabilizing effect.

Example 6 (Welfare Reduction via Cost Increase; Haller (2012), Example 5)

Let $N = \{1, 2, 3\}$, $V_{ij} = V = 24$ for all $i \neq j$, $c_{13} = 32, c_{23} = 28$, $c_{12} = 56, c_{32} = 40$, and $c_{21} = c_{31} = 80$.

The periphery-sponsored star with center 3 is the only (strict) Nash network in the absence of a publicly provided infrastructure. Now suppose that the publicly provided infrastructure $\mathfrak{g} = \{12\}$ is installed. Then $g' = \{32\}$ is the only (strict) Nash network. The corresponding outcome is $\mathfrak{g} \oplus \{32\}$. In $\mathfrak{g} \oplus \{32\}$, each of the two links costs more than a link in the center-sponsored star with center 3. $\square$

This example is reminiscent of Braess's paradox [Braess (1968), Braess *et al.* (2005)] that an extension of a road network may cause a redistribution of traffic which results in longer individual travel times. In the present example, installation of the public infrastructure causes players to switch to costlier links.

6 On the Provision of Public Infrastructure

We have seen that the public provision of infrastructure followed by decentralized network formation can be beneficial or undesirable: it can be stabilizing or destabilizing, respectively; welfare improving or welfare reducing, respectively. In some instances, a pre-existing infrastructure ought to be removed or not to be used any longer. If the latter can be achieved at no cost, say by just leaving a cable in the ground, then a benevolent, omniscient, and omnipotent policy maker may do. Otherwise, the abandonment, removal, renewal, expansion, or replacement of a publicly provided infrastructure can be a formidable political challenge.

In the remainder of this section, I shall focus on the simpler task of creating a public infrastructure de novo, with the absence of a public infrastructure as the status quo. Policy makers as well as players compare a Nash network g^* of the model without a public infrastructure with an outcome $\mathfrak{g} \oplus g^{**}$ where $\mathfrak{g} \neq e$ is a candidate for a publicly provided infrastructure and g^{**} is a Nash network of the strategic game of network formation with payoff functions $\Pi_i(\mathfrak{g}; g)$, $g \in \mathcal{G}$, $i \in N$. To simplify the analytic task even further, I assume that

(i) the choice is between the empty infrastructure e and the proposed infrastructure project $\mathfrak{g} \neq e$;

(ii) all players anticipate the Nash equilibrium g^* in the subsequent network formation game if the public infrastructure e is in place;

(iii) all players anticipate the Nash equilibrium g^{**} in the subsequent network formation game if the public infrastructure $\mathfrak{g}$ is in place.

Then frequently, a version of the well known "subscription game" will yield selection and sufficient funding of the more efficient of the two infrastructures. In that game, the infrastructure $\mathfrak{g}$ gets built if the sum of the players' voluntary contributions s_i, $i \in N$, equals or exceeds the cost of constructing $\mathfrak{g}$. In that case, the players forfeit any contributions in excess of the construction cost. The infrastructure $\mathfrak{g}$ is not built, if the aggregate voluntary contributions fall short of the construction cost. In that case, the contributions are refunded. The actions in the following strategic game Γ represent the individual contributions s_i, $i \in N$, and the payoffs reflect the players' evaluations of the resulting outcomes in accordance with (i)–(iii). Let $\Gamma = (N, (S_i)_{i \in N}, (U_i)_{i \in N})$ where

$$\begin{array}{ll} s_i \in S_i = \mathbb{R}_+ \text{ for } i \in N; & \\ U_i(s_1, \ldots, s_n) = \Pi_i(g^*) & \text{if } \sum_{j \in N} s_j < \mathfrak{c}(\mathfrak{g}); \\ U_i(s_1, \ldots, s_n) = \Pi_i(\mathfrak{g}; g^{**}) - s_i & \text{if } \sum_{j \in N} s_j \geq \mathfrak{c}(\mathfrak{g}); \\ \mathfrak{c}(\mathfrak{g}) = \sum_{j \neq k} \mathfrak{g}_{jk} \cdot c_{jk}, & \text{is the cost of constructing } \mathfrak{g}. \end{array}$$

To analyze Γ, set $\Delta_i = \Pi_i(\mathfrak{g}; g^{**}) - \Pi_i(g^*)$ for $i \in N$. Δ_i is i's equilibrium payoff differential between the ensuing network formation games $(N, (\mathcal{G}_i)_{i \in N}, (\Pi_i(\mathfrak{g}; \cdot))_{i \in N})$ and $(N, (\mathcal{G}_i)_{i \in N}, (\Pi_i)_{i \in N})$ when the Nash equilibria g^* and g^{**} obtain. Further set $\Delta = \sum_i \Delta_i$, $N_+ = \{i \in N : \Delta_i > 0\}$, and $\Delta_+ = \sum_{i \in N_+} \Delta_i \geq \Delta$. Then

$$W_0(\mathfrak{g} \oplus g^{**}) - W_0(g^*) = \Delta - \mathfrak{c}(\mathfrak{g}). \tag{4}$$

On efficiency grounds, $\mathfrak{g}$ should be built if $W_0(\mathfrak{g} \oplus g^{**}) > W_0(g^*)$ or equivalently, by (4), if $\Delta > \mathfrak{c}(\mathfrak{g})$. And $\mathfrak{g}$ should not be built if $W_0(\mathfrak{g} \oplus g^{**}) < W_0(g^*)$ (or equivalently, $\Delta < \mathfrak{c}(\mathfrak{g})$).

Proposition 3

(a) *Suppose* $W_0(\mathfrak{g} \oplus g^{**}) > W_0(g^*)$.

(aa) *A strict Nash equilibrium of* Γ *exists.*

(ab) $\sum_i s_i^* = \mathfrak{c}(\mathfrak{g})$ *in every strict Nash equilibrium* $(s_1^*, \ldots, s_n^*)$ *of* Γ.

(b) *Suppose* $W_0(\mathfrak{g} \oplus g^{**}) < W_0(g^*)$.

(ba) *If* $\Delta_i < \mathfrak{c}(\mathfrak{g})$ *for all* i, *then a strict Nash equilibrium of* Γ *exists, namely* $(s_1^*, \ldots, s_n^*) = (0, \ldots, 0)$.

(bb) *If* $\Delta_i \geq 0$ *for all* i, *then* $\sum_i s_i^* < \mathfrak{c}(\mathfrak{g})$ *in every Nash equilibrium* $(s_1^*, \ldots, s_n^*)$ *of* Γ.

(bc) *If* $\Delta_+ > \mathfrak{c}(\mathfrak{g})$, *then there exists a strict Nash equilibrium* $(s_1, \ldots, s_n)$ *of* Γ *with* $\sum_i s_i = \mathfrak{c}(\mathfrak{g})$.

PROOF. (aa) Suppose $W_0(\mathfrak{g} \oplus g^{**}) > W_0(g^*)$. Then $\Delta > \mathfrak{c}(\mathfrak{g})$ and, consequently, $\Delta_+ \geq \Delta > \mathfrak{c}(\mathfrak{g}) > 0$. Set $s_i = (\mathfrak{c}(\mathfrak{g})/\Delta_+) \cdot \Delta_i$ for $i \in N_+$ and $s_i = 0$ for $i \notin N_+$. Then $s_i \in S_i$ for all i and $\sum_i s_i = \mathfrak{c}(\mathfrak{g})$.

For $i \notin N_+$, $U_i(s_1, \ldots, s_n) = \Pi_i(\mathfrak{g}; g^{**})$ and $s_i = 0$. Choosing $s_i' > 0$ given s_{-i} would yield payoff $\Pi_i(\mathfrak{g}; g^{**}) - s_i' < \Pi_i(\mathfrak{g}; g^{**})$. Hence s_i is i's unique best response against s_{-i}.

For $i \in N_+$, $U_i(s_1, \ldots, s_n) = \Pi_i(\mathfrak{g}; g^{**}) - s_i$ and $0 < s_i = (\mathfrak{c}(\mathfrak{g})/\Delta_+) \cdot \Delta_i < \Delta_i$. Choosing $s_i' > s_i$ given s_{-i} would yield payoff $\Pi_i(\mathfrak{g}; g^{**}) - s_i' < \Pi_i(\mathfrak{g}; g^{**}) - s_i$. Choosing $s_i' < s_i$ given s_{-i} would yield payoff $\Pi_i(g^*) = \Pi_i(\mathfrak{g}; g^{**}) - \Delta_i < \Pi_i(\mathfrak{g}; g^{**}) - s_i$. Hence s_i is i's unique best response against s_{-i}. This shows that $(s_1, \ldots, s_n)$ is a strict Nash equilibrium of Γ.

(ab) Suppose $W_0(\mathfrak{g} \oplus g^{**}) > W_0(g^*)$. Now let $(s_1^*, \ldots, s_n^*)$ be a strict Nash equilibrium of Γ. Then $\sum_i s_i^* = \mathfrak{c}(\mathfrak{g})$. Suppose not. If $\sum_i s_i^* > \mathfrak{c}(\mathfrak{g})$, then $U_i(s_1^*, \ldots, s_n^*) = \Pi_i(\mathfrak{g}; g^{**}) - s_i^*$ for all i and $s_j^* > 0$ for some j. If such a player reduces his contribution by a sufficiently small amount $\varepsilon_j > 0$, given s_{-j}^*, then $s_j^* - \varepsilon_j \geq 0$, $s_j^* - \varepsilon_j + \sum_{i \neq j} s_i^* \geq \mathfrak{c}(\mathfrak{g})$ and $U_j(s_j^* - \varepsilon_j; s_j^*) = \Pi_i(\mathfrak{g}; g^{**}) - (s_i^* - \varepsilon_j) > \Pi_i(\mathfrak{g}; g^{**}) - s_i^* = U_j(s_1^*, \ldots, s_n^*)$, contradicting the assumption that $(s_1^*, \ldots, s_n^*)$ is a Nash equilibrium of Γ. If $\sum_i s_i^* < \mathfrak{c}(\mathfrak{g})$, then $U_i(s_1^*, \ldots, s_n^*) = \Pi_i(g^*)$ for all i. If some player j increases his contribution by a sufficiently small amount $\varepsilon_j > 0$, given s_{-j}^*, then $s_j^* + \varepsilon_j \geq 0$, $s_j^* + \varepsilon_j + \sum_{i \neq j} s_i^* < \mathfrak{c}(\mathfrak{g})$ and $U_j(s_j^* - \varepsilon_j; s_j^*) = \Pi_j(g^*) = U_j(s_1^*, \ldots, s_n^*)$, contradicting the assumption that $(s_1^*, \ldots, s_n^*)$ is a strict Nash equilibrium of Γ. Thus $\sum_i s_i^* \neq \mathfrak{c}(\mathfrak{g})$ always leads to a contradiction. Hence as asserted, $\sum_i s_i^* = \mathfrak{c}(\mathfrak{g})$ has to hold.

(ba) Suppose $W_0(\mathfrak{g} \oplus g^{**}) < W_0(g^*)$ and $\Delta_i < \mathfrak{c}(\mathfrak{g})$ for all i. Consider $(s_1, \ldots, s_n) = (0, \ldots, 0)$. Then $\sum_i s_i = 0 < \mathfrak{c}(\mathfrak{g})$ and $U_i(s_1, \ldots, s_n) = \Pi_i(g^*)$ for all i. If some player j chooses $s'_j \in (0, \mathfrak{c}(\mathfrak{g}))$ given s_{-j}, then $s'_j + \sum_{i \neq j} s_i = s'_j < \mathfrak{c}(\mathfrak{g})$ and $U_j(s'_j, s_{-j}) = \Pi_i(g^*) = U_j(s_1, \ldots, s_n)$. If j chooses $s'_j \geq \mathfrak{c}(\mathfrak{g})$ given s_{-j}, then $s'_j + \sum_{i \neq j} s_i = s'_j \geq \mathfrak{c}(\mathfrak{g})$ and $U_j(s'_j, s_{-j}) = \Pi_j(\mathfrak{g}; g^{**}) - s'_j \leq \Pi_j(\mathfrak{g}; g^{**}) - \mathfrak{c}(\mathfrak{g}) < \Pi_j(\mathfrak{g}; g^{**}) - \Delta_j = \Pi_i(g^*) = U_j(s_1, \ldots, s_n)$. Hence s_j is a best response against s_{-j} for all j. This shows that $(0, \ldots, 0)$ is a Nash equilibrium of Γ.

(bb) Suppose $W_0(\mathfrak{g} \oplus g^{**}) < W_0(g^*)$ and $\Delta_i \geq 0$ for all i. Now let $(s_1^*, \ldots, s_n^*)$ be a Nash equilibrium of Γ. Suppose $\sum_i s_i^* \geq \mathfrak{c}(\mathfrak{g})$. Then $U_i(s_1^*, \ldots, s_n^*) = \Pi_i(\mathfrak{g}; g^{**}) - s_i^*$ for all i. By the same argument as in the proof of (ab), $\sum_i s_i^* > \mathfrak{c}(\mathfrak{g})$ can be ruled out. Further, $\Pi_i(g^*) = \Pi_i(\mathfrak{g}; g^{**}) - \Delta_i$ for each player i. Hence $s_i^* \leq \Delta_i$ has to hold for $(s_1^*, \ldots, s_n^*)$ to be a Nash equilibrium of Γ with $\sum_i s_i^* = \mathfrak{c}(\mathfrak{g})$. But then $\sum_i s_i^* \leq \sum_i \Delta_i = \Delta < \mathfrak{c}(\mathfrak{g})$ because of $W_0(\mathfrak{g} \oplus g^{**}) < W_0(g^*)$, in contradiction to $\sum_i s_i^* = \mathfrak{c}(\mathfrak{g})$. Hence to the contrary, $\sum_i s_i^* < \mathfrak{c}(\mathfrak{g})$ has to hold.

(bc) Suppose $\Delta_+ > \mathfrak{c}(\mathfrak{g})$. Then a strict Nash equilibrium $(s_1, \ldots, s_n)$ of Γ with $\sum_i s_i = \mathfrak{c}(\mathfrak{g})$ can be constructed as in the proof of (aa). ■

Proposition 3 warrants several comments and qualifications. First of all, if costs are owner-homogeneous, an equilibrium outcome of Γ, e or $\mathfrak{g}$, defines a subgame-perfect equilibrium outcome g^* or $\mathfrak{g} \oplus g^{**}$, respectively, of the two-stage game where Γ is played in the first stage and depending on the outcome of the first stage, $(N, (\mathcal{G}_i)_{i \in N}, (\Pi_i)_{i \in N})$ or $(N, (\mathcal{G}_i)_{i \in N}, (\Pi_i(\mathfrak{g}; \cdot))_{i \in N})$ is played in the second stage.[4] A second-stage game may have multiple equilibria. To obtain a well defined subscription game, I assume that the players expect g^* to prevail in $(N, (\mathcal{G}_i)_{i \in N}, (\Pi_i)_{i \in N})$ and $\mathfrak{g} \oplus g^{**}$ to prevail in $(N, (\mathcal{G}_i)_{i \in N}, (\Pi_i(\mathfrak{g}; \cdot))_{i \in N})$. A similar dilemma of multiple second-stage equilibria — that needs to be resolved by specific expectations — arises in two-stage resource allocation mechanism where the first stage is under the planner's control whereas the second stage consists of competitive exchange outside the planner's control. See Guesnerie (1995), Section 3.1.

[4]By Proposition 2, owner-homogeneity of costs is sufficient for existence of second-stage equilibria. Without owner-homogeneity of costs, both existence and non-existence can occur.

Further Remarks

1. The restriction to strict Nash equilibria in Proposition 3(ab) is crucial. Take for instance the backbone infrastructure $\mathfrak{g} = \{48\}$ in the following adaptation of Example 3 in Haller (2012) where such a publicly provided link — which would not be created by any player — does not alter the set of second-stage Nash equilibria, yet constitutes a strict Pareto improvement:

Example 7 (Strict Pareto Improvement)
Let $N = \{1, \ldots, 8\}$, $K = \{1, 2, 3, 4\}$, $L = \{5, 6, 7, 8\}$. Further set for $i \neq j$, $V_{ij} = 1$ and

$$c_{ij} = \begin{cases} 0.8 & \text{if } i \neq j,\ i, j \in K \text{ or } i, j \in L; \\ 16 & \text{otherwise.} \end{cases}$$

In the absence of a publicly provided infrastructure, i.e., in case $\mathfrak{g} = e$, a Nash network g^* has two minimally connected components, one consisting of the members of K and the other one consisting of the members of L.

Next suppose that the link $ij = 48$ is publicly provided, that is $\mathfrak{g} = \{48\}$. Then the set of Nash equilibria remains the same, but the total benefit is increased by $8 \times 4 = 32$, at a cost of 16 for the link 48. This constitutes a strict Pareto improvement if there is a way to charge each player a cost contribution of 2 for the link 48.

Now take the backbone infrastructure $\mathfrak{g} = \{48\}$. We may choose $g^* = g^{**}$. Then $W_0(\mathfrak{g} \oplus g^{**}) > W_0(g^*)$. But $(0, \ldots, 0)$ is a Nash equilibrium of Γ which is not strict. $\square$

2. Under the assumptions of Proposition 3(bb), strict Nash equilibria do not exist and, therefore, restricting attention to strict Nash equilibria proves counter-productive.

3. The outcome in Proposition 3(bc) is the public provision of the infrastructure $\mathfrak{g}$ even though aggregate welfare is diminished after the creation of the infrastructure. The beneficiaries from $\mathfrak{g}$, the members of N_+, are willing to shoulder the entire cost of construction while others are opposed to the creation of $\mathfrak{g}$. Such a scenario is conceivable, both in practice and in theory. In Example 6, consider $\mathfrak{g} = \{12\}$, $g^* = \{13, 23\}$ (the periphery-sponsored star with center 3), and $g^{**} = \{32\}$ so that $\mathfrak{g} \oplus g^{**}$ is the periphery-sponsored star with center 2. Then $\mathfrak{c}(\mathfrak{g}) = 56$, $\Delta = 20$, and $\Delta_+ = 60$, with $N_+ = \{1, 2\}$. Hence the assumptions of Proposition 3(bc) are met.

4. The general approach taken in this section and the equilibrium analysis in Proposition 3 share the basic features of the model of Bagnoli and Lipman (1989), but differ in many details, in particular equilibrium selection. Here we deal with the special case $M = 1$ in Bagnoli and Lipman (1989). Bagnoli and McKee (2001) provide episodical empirical support as well as experimental support for that case. $M = 1$ in our context means a choice between e and $\mathfrak{g} \neq e$. $M > 1$ would mean to choose an infrastructure from among $M + 1$ different infrastructures e, $\mathfrak{g}(1), \mathfrak{g}(2), \ldots, \mathfrak{g}(M)$. The corresponding analysis of Bagnoli and Lipman (1989) shows that this case poses a more intricate mechanism design problem, which I leave to future research.

5. The "subscription model" or "subscription game" has been investigated numerous times, under a variety of names, both theoretically (most notably by Bergstrom, Blume, and Varian (1986)) and experimentally (by Andreoni (1988) and many others). It has become part of the textbook literature in public economics, e.g., Cornes and Sandler (1986) and Myles (1995). In the standard model, the aggregate contribution $\sum_i s_i$ determines the amount q of a public good: $q = \sum_i s_i$ or, more generally, $q = f(\sum_i s_i)$. It is a widely held wisdom among economists that the non-cooperative equilibrium outcome of the model is inefficient and in fact, there is under-provision of the public good. It was the basic insight of Bagnoli and Lipman and others that efficient equilibrium outcomes could obtain if the problem was selection of a discrete public project from a finite set rather than choosing an amount q from a continuum of possible amounts. With the exception of conclusion (bc), Proposition 3 further confirms their insight.

6. The "voluntary contributions game" of Section 4 is an incarnation of the "subscription model". Yet the "fixed fee game" of Section 4 is not a subscription game à la Bagnoli and Lipman because the amount of public good varies with the number of players paying the fee and players cannot choose arbitrary non-negative payments.

7. The idea to finance the cost of building a network via voluntary contributions also appears in Anshelevich *et al.* (2008). In their network formation game, each player i decides on her non-negative contribution $p_i([jk])$ to the cost $c_{[jk]}$ of each potential edge $[jk]$. Edge $[jk]$ is created if $\sum_i p_i([jk]) \geq c_{[jk]}$. In Bloch and Jackson (2007), players can subsidize other players to form specific links and can also bribe other players not to form particular links.

8. One could in principle start with an efficient public infrastructure $\mathfrak{g}$, which would imply $\Delta \geq \mathfrak{c}(\mathfrak{g})$ and existence of at least one (not necessarily strict) Nash equilibrium $(s_1, \ldots, s_n)$ of Γ with $\sum_i s_i \geq \mathfrak{c}(\mathfrak{g})$. By Proposition

2 of Haller (2012), however, efficiency of $\mathfrak{g}$ renders the private provision of a network g built around or upon $\mathfrak{g}$ obsolete. In addition, distrust and disbelief in big government could preempt huge public projects. For instance, in the insider-outsider model of Section 7 in Haller (2012), public provision of a backbone infrastructure might be acceptable while public provision of a larger infrastructure might not. Notice that a backbone infrastructure $\mathfrak{g}$ by itself is inefficient, but may lead to an efficient equilibrium outcome $\mathfrak{g} \oplus g^{**}$. There could also exist an economic reason — which would warrant a slight modification of the model — why very expensive public infrastructures should be avoided. A conceivable reason could be a "shadow cost of public funds" $\lambda > 0$ which could mean, for example, that 'distortionary taxation inflicts disutility \$$(1+\lambda)$ on taxpayers in order to levy \$1 for the state.'[5] In the present context, λ may simply account for administrative or overhead costs. Needless to say, the private sector may incur overhead costs as well, but possibly less than the public sector.

9. Mutuswami and Winter (2002) address the question "how a social planner can ensure the formation of an efficient network in a scenario where the costs of network formation are publicly known but an individual player's benefits from network formation are not known to him." They consider mechanisms where a planner decides upon the network and cost contributions, based on the desired links and cost contributions announced by the agents. They show that a mechanism can be designed that meets three criteria: efficiency, balanced budget, and equity. Their approach as it stands yields an efficient publicly provided network and (in view of Proposition 2 of Haller (2012)) leaves no room for decentralized network formation.

7 Final Remarks

The paper's basic analysis rests on two main premises: (a) A network is regarded as a public infrastructure. (b) The network or part of it is the outcome of a game in strategic form. The basic investigation is devoted to efficiency criteria for networks and differences and parallels between strategic network formation and funding schemes for public goods.

The analysis of Sections 5 and 6 adds two more premises: First, strategic network formation builds a network around or upon a core network or infrastructure. Second, the core network is either pre-existing or publicly provided

[5] Laffont and Tirole (1993, p. 55).

prior to the onset of strategic network formation. Section 5 reports on prior findings of Haller (2012), in particulars those related to efficiency. Section 6 is devoted to the choice and funding of a publicly provided infrastructure. The problem is confined to a binary choice between the absence of a public infrastructure (choice of the empty network e) and the public provision of a particular infrastructure $\mathfrak{g}$ at cost $\mathfrak{c}(\mathfrak{g})$. Everybody anticipates that the choice of e will lead to a specific Nash equilibrium outcome g^* in the subsequent network formation game whereas the choice of $\mathfrak{g}$ will yield the Nash equilibrium outcome $\mathfrak{g} \oplus g^{**}$. It turns out that the more efficient alternative is frequently chosen in an equilibrium of a suitable version of the subscription game.

The foregoing analysis, based on the connections model with two-way flow of information (without decay and with perfectly reliably links), presumes positive network externalities. Haller (2012) demonstrates by way of two examples that some of his analysis can be extended to instances with **negative network externalities** of the following form: The values V_{ij} are endogenously determined and of the form $V_{ij}(\overline{g}) \equiv v_{ij}/d_j(\overline{g})$ where for $j \in N, g \in \mathcal{G}$, $d_j(\overline{g}) = \sum_{k \neq j} \overline{g}_{jk}$ denotes the number of j's direct neighbors in $\overline{g}$, the degree of node j in $\overline{g}$. The idea is that the information originating from a player j is harder to access if j has more direct neighbors. More specifically, there are exogenously given undiluted values v_{ij} for $i \neq j$. If $j \in N_i(\overline{g})$, then $d_j(\overline{g}) \geq 1$ and $V_{ij}(\overline{g}) \equiv v_{ij}/d_j(\overline{g})$. Möhlmeier et al. (2013) assume cost and value homogeneity, but allow for decay. Their general formulation is non-parametric. In their parametric examples, they express negative externalities by means of coefficients $1/(1 + d_j(g))$ rather than $1/d_j(\overline{g})$. Subsection 4.3 in Haller and Sarangi (2005) considers negative link externalities as well and pursues a similar idea: A node might become less reliable, less effective in providing information via its direct links, if it gets accessed through one more direct link. Morrill (2011) considers "degree-based utility functions" where a player only receives benefits $\phi(d_j(g))$ from direct neighbors j, with $\phi(d_j(g))$ decreasing in $d_j(g)$. The coauthor model of Jackson and Wolinsky (1996) and networks of buyers and sellers à la Kranton and Minehart (2001) share that feature. Goyal and Joshi (2006) investigate the effect of positive and negative externalities on network formation. Buechel and Hellmann (2012) study the relationship between the two kinds of network externalities and over- or under-connectivity of stable networks.

Durieu, Haller, and Solal (2011) allow for the possibility that a person's increased networking efforts not only establish or strengthen desirable links to specific agents, but also cause more calls from phone banks, more spam, more encounters with annoying or hostile people. They present an example where positive and negative externalities coexist and refraining from networking is a strictly dominant strategy while it would be efficient if everybody networked. From the initiating individual's perspective, the negative effects of a link dominate whereas from a social perspective the net effect is positive. The role of negative network externalities and the interaction between positive and negative externalities deserve further scrutiny.

References

Andreoni, J. (1988), "Why Free Ride?", *Journal of Public Economics*, 37, 291-304.

Anshelevich, E., Dasgupta, A., Tardos, É., and T. Wexler (2008), "Near-Optimal Network Design with Selfish Agents," *Theory of Computing*, 4, 77-109.

Bala, V. and S. Goyal (2000), "A Non-Cooperative Model of Network Formation", *Econometrica,* 68, 1181-1229.

Bagnoli, M. and B.L. Lipman (1989), "Limited Provision of Public Goods: Fully Implementing the Core through Private Contributions," *Review of Economic Studies*, 56, 583-601.

Bagnoli, M. and M. McKee (1991), "Voluntary Contribution Games: Efficient Private Provision of Public Goods," *Economic Inquiry*, 29, 351-366.

Bergstrom, T., Blume, L. and H. Varian (1986), "On the Private Provision of Public Goods," *Journal of Public Economics*, 29, 25-49.

Billand, P., Bravard, C. and S. Sarangi (2008), "Existence of Nash Networks in One-Way Flow Models," *Economic Theory*, 37, 491-507.

Bloch, F. and M.O. Jackson (2007), "The Formation of Networks with Transfers among Players," *Journal of Economic Theory*, 133, 83-110.

Braess, D. (1968), "Über ein Paradoxon aus der Verkehrsplanung," *Unternehmensforschung*, 12, 258-268.

Braess, D., Nagurney, A. and T. Wakolbinger (2005), "On a Paradox of Traffic Planning," *Transportation Science*, 39, 446-450.

Buechel, B. and T. Hellmann (2012), "Under-connected and Over-connected Networks: the Role of Externalities in Strategic Network Formation," *Review of Economic Design*, 16, 71-87.

Cornes, R. and T. Sandler (1986), *The Theory of Externalities, Public Goods, and Club Goods*. Cambridge University Press: Cambridge, UK.

Derks, J. and M. Tennekes (2009), "A Note on the Existence of Nash Networks in One-Way Flow Models," *Economic Theory*, 41, 515-522.

Durieu, J., Haller, H. and P. Solal (2011), "Nonspecific Networking," *Games*, 2, 87-113. Accessible at http://www.mdpi.com/2073-4336/2/1/87/pdf.

Gallager, R.G., Humblet, P.A. and P.M. Spira (1983), "A Distributed Algorithm for Spanning Trees," *ACM Transactions on Programming Languages and Systems*, 5, 66-77.

Galeotti, A. (2006), "One-Way Flow Networks: the Role of Heterogeneity," *Economic Theory*, 29, 163-179.

Galeotti, A., Goyal, S. and J. Kamphorst (2006), "Network Formation with Heterogeneous Players," *Games and Economic Behavior,* 54, 353-372.

Goyal, S. and S. Joshi (2006), "Unequal Connections," *International Journal of Game Theory,* 34, 319-349.

Guesnerie, R. (1995), *A Contribution to the Pure Theory of Taxation.* Cambridge University Press: Cambridge, UK.

Haller, H. (2012), "Network Extension," *Mathematical Social Sciences*, 64, 166-172.

Haller, H., Kamphorst, J. and S. Sarangi (2007), "(Non-)Existence and Scope of Nash Networks," *Economic Theory*, 31, 597-604.

Haller, H. and S. Sarangi (2005), "Nash Networks with Heterogeneous Links," *Mathematical Social Sciences*, 50, 181-201.

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

Kranton R. and D. Minehart (2001), "A Theory of BuyerSeller Networks," *American Economic Review*, 91, 485-508.

Laffont, J.-J. and J. Tirole (1993), *A Theory of Incentives in Procurement and Regulation*. The MIT Press: Cambridge, MA.

Morrill, T. (2011), "Network Formation under Negative Degree-based Externalities," *International Journal of Game Theory*, 40, 367-385.

Möhlmeier, P., Rusinowska, A. and E. Tanimura (2013), "A Degree-Distance-based Connections Model with Negative and Positive Externalities," Centre d'Economie de la Sorbonne, Université Paris 1, Working Paper 2013.40.

Mutuswami, S. and E. Winter (2002), "Subscription Mechanisms for Network Formation," Journal of Economic Theory, 106, 242-264.

Myerson, R.B. (1991), *Game Theory: Analysis of Conflict*. Harvard University Press: Cambridge, MA.

Myles, G.D. (1995), *Public Economics*. Cambridge University Press: Cambridge, UK.