Bonn Econ Discussion Papers

Discussion Paper 28/2005

Network Potentials

by

Subhadip Chakrabarti, Robert P. Gilles

October 2005

Bonn Graduate School of Economics
Department of Economics
University of Bonn
Adenauerallee 24 - 42
D-53113 Bonn

The Bonn Graduate School of Economics is
sponsored by the

Deutsche Post World Net

MAIL EXPRESS LOGISTICS FINANCE

Network Potentials*

Subhadip Chakrabarti† Robert P. Gilles‡

October 2005

Abstract

A network payoff function assigns a utility to all participants in a (social) network. In this paper we discuss properties of such network payoff functions that guarantee the existence of certain types of pairwise stable networks and the convergence of certain network formation processes. In particular we investigate network payoff functions that admit an exact network potential or an ordinal network potential. We relate these network potentials to exact and ordinal potentials of a non-cooperative network formation game based on consent in link formation. Our main results extend and strengthen the current insights in the literature on game theoretic approaches to social network formation.

Keywords: Network formation; pairwise stability; potential functions.

JEL Classification: C72, C79, D85.

---

*We thank Matt Jackson and Sudipta Sarangi for extensive discussions on the subject of this paper. For downloadable PDF files of several of the cited papers we refer to the corresponding author's web site at `http://gilles.economics-network.org`.

†Department of Economics, Queen's University, Belfast, Northern Ireland, UK. *Email*: schakrab@vt.edu. Part of this research was done while this author was at Bonn on a post-doctoral research fellowship. We thank the Department of Economics at the University of Bonn for their hospitality and financial support.

‡**Corresponding author.** *Address*: Department of Economics, Virginia Tech (0316), Blacksburg, VA 24061, USA. *Email*: rgilles@vt.edu. Part of this research was done at the Center for Economic Research at Tilburg University, Tilburg, the Netherlands. Financial support from the Netherlands Organization for Scientific Research (NWO), grant # 46-550, is gratefully acknowledged.

1 Introduction

Since the introduction of the notion of pairwise stability in Jackson and Wolinsky (1996), one of the main research questions that have been investigated in the literature on social network formation, has been the problem of the existence of (strictly) pairwise stable networks. A related question is whether certain network formation processes converge to these types of stable networks. Preliminary investigation of these issues was addressed in Jackson and Watts (2001) and Jackson and Watts (2002b). Here we investigate the fundamental properties of network payoff structures that bear on these two fundamental problems. We show that the admittance of (ordinal) network potential functions provides a powerful tool that satisfactorily resolves these fundamental problems.

We consider undirected social networks, which are founded on the principle that each link is between two equally participating players. Hence, both players are assumed to consent voluntarily and in full knowledge to the formation of a link between them.[1] This idea was first formalized by Myerson (1991, page 448) as a normal form game—referred to as the "Consent Game". Myerson's Consent Game and its variations unfortunately admit a multitude of Nash equilibria as pointed out in Dutta, van den Nouweland, and Tijs (1998), Jackson (2005b), and Gilles, Chakrabarti, and Sarangi (2005a). From this rather unsatisfactory fact, there emerged two alternative theories of network formation under consent.

In their seminal paper Jackson and Wolinsky (1996) introduced a *link-based* approach to network formation. The central strategic concept is now the link rather than the players between whom this link is formed. In particular, they defined the link-based notion of *pairwise stability* as an alternative description of the requirement of consent in link formation. This alternative formulation of stability reduces the multitude of acceptable networks. Existence of pairwise stable networks and the convergence of simple link-based improvement processes to such pairwise stable networks has been addressed in Jackson and Watts (2001), Jackson and Watts (2002a), and Jackson and Watts (2002b).

The other approach to resolve the problem of the indeterminacy of Myerson's Consent Game is to consider *network-based refinements* of the Nash equilibrium concepts within the context of Myerson's Consent Game and its variations. This has been pursued by Gilles, Chakrabarti, and Sarangi (2005a) and Gilles and Sarangi (2005a)

[1] The alternative is to consider directed social networks, in which links are formed without the consent of one of the players. We refer to Bala and Goyal (2000) and related subsequent work for a study of these so-called *Nash networks*.

for the case that the formation of links is costly. Gilles and Sarangi (2005a) consider myopic learning processes based on the formation of simple belief systems. In these belief systems a player formulates expectations about whether other players would consent to forming links with her. Each player now formulates a best response to her formed myopic beliefs. This learning process only converges satisfactorily if the formed beliefs are confirmed within the stable state that is reached. Gilles and Sarangi show that the resulting set of these so-called self-confirming equilibria generate a strict subset of the class of pairwise stable networks, the so-called *strictly pairwise stable networks*. Gilles and Sarangi (2005a) do not address under which conditions such strictly pairwise stable networks exist.

Here we show that if the network payoff function admits a potential function, the existence of such strictly pairwise stable networks is guaranteed and, therefore, that the learning processes based on the formation of the myopic belief systems introduced by Gilles and Sarangi (2005a) indeed converge to self-confirming equilibrium networks. Furthermore we explore the relationship of these network potentials with the improvement processes discussed by Jackson and Watts (2002a).

Two seminal contributions to the game theoretic literature address the introduction of potential functions into the analysis of cooperative as well as non-cooperative games. These potential functions were seminally introduced by Hart and Mas-Colell (1989) for cooperative games and by Monderer and Shapley (1996) for non-cooperative games. Hart and Mas-Colell showed the fundamental relationship between their cooperative potential function and the Shapley value for cooperative games. (Shapley 1953) Monderer and Shapley and subsequent contributions showed that non-cooperative games admitting potentials are closely related to congestion games in the sense of Rosenthal (1973). These two strands of the literature were brought together by Qin (1996) and especially Ui (2000), who showed that the payoff functions in non-cooperative potential games are the Shapley values of a corresponding class of cooperative games. These fundamental insights showed that there is a strong relationship between the various classes of non-cooperative potential games and Hart and Mas-Colell's cooperative potential theory.

In the present paper we introduce potential functions for a broad class of network payoff functions as descriptors of network formation fundamentals. First, we consider *exact* network potentials, which assign the marginal contributions of newly formed links to the participants in the network. The class of network payoff functions for which such exact network potentials can be considered are exactly those network payoff functions that satisfy Myerson (1977)'s "fairness" property—more recently reformulated as an *equal bargaining power* property. (Jackson and Wolinsky 1996, page

54) We show that a network payoff function admits an exact network potential if and only if Myerson's Consent Game admits a game-theoretic exact potential in the sense of Monderer and Shapley (1996). Furthermore, we extend Qin (1996)'s equivalence theorem by showing that, if a network payoff function admits such an exact network potential, the network payoffs correspond in fact to the Shapley values of a certain class of related cooperative games.

The presence of an exact network potential is a rather demanding property. It is natural to weaken this requirement to that of the admittance of an *ordinal* network potential function. We show that the presence of an ordinal network potential is a strictly weaker requirement than Myerson's Consent Game admitting a game-theoretic ordinal potential in the sense of Monderer and Shapley (1996).

The main consequence of the presence of an ordinal potential is to guarantee the convergence of network formation processes based on myopic payoff improvements as introduced in Jackson and Watts (2002a) and the existence of strictly pairwise stable networks. In particular, if the network payoff function admits an ordinal network potential, such network formation processes based on myopic payoff improvements are shown to converge to pairwise stable networks. This generalizes the main result in Jackson and Watts (2001). Furthermore, if Myerson's Consent Game admits an ordinal potential, there exists at least one strictly pairwise stable network in the sense of Gilles and Sarangi (2005a).

It should be clear from the developed insights in this paper that the concept of a network potential is a very powerful one. It allows the two main questions regarding the existence of (strictly) pairwise stable networks and the convergence of certain network formation processes to be fully resolved.

The rest of the paper is organized as follows. Section 2 introduces terminology, the notation and the basic network formation theories. In section 3, we define exact network potentials and their relation with the exact potentials of the Consent Game introduced by Myerson (1991, page 448). In Section 4, we define ordinal network potentials and their relationship with ordinal potentials of Myerson's Consent Game. In Section 5, we show that certain classes of network formation games admitting (ordinal) potentials allow the convergence of network formation processes to pairwise stable networks and guarantee the existence of a strictly pairwise stable network, thus linking this research to work on network formation based on myopic belief systems. (Gilles and Sarangi 2005a) In Section 6 we look at some applications of potential functions and some special properties of these classes of network formation games. Section 7 concludes.

2 Preliminaries

In this section we discuss required material from non-cooperative game, cooperative game, the theory of game theoretic potentials, and network formation theory.

2.1 Non-cooperative games and potentials

A *non-cooperative game* on a fixed, finite player set $N = \{1, \ldots, n\}$ is given by a list $(A, \pi) = (A_i, \pi_i)_{i \in N}$ where for every player $i \in N$, A_i denotes an action set and $\pi_i : A \to \mathbb{R}$ denotes player i's payoff function with $A = A_1 \times A_2 \times \cdots \times A_n$. An individual action of player $i \in N$ is denoted by $a_i \in A_i$ and an action tuple is written as $a = (a_1, a_2, \ldots, a_n) \in A$. For every action tuple $a \in A$ and player $i \in N$, we denote by $a_{-i} = (a_1, a_2, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n) \in A_{-i} = \prod_{j \neq i} A_j$ as the actions selected by players other than i. $\pi = (\pi_1, \pi_2, \ldots, \pi_n) : A \to \mathbb{R}^N$ is the composite payoff function of this game.

Throughout this paper, we denote a non-cooperative game for short by the pair (A, π). A non-cooperative game is finite if for every $i \in N$, the action set A_i is finite. In this paper, we only apply finite games.

An action set $a_i \in A_i$ for player $i \in N$ is a *best response* to $a_{-i} \in A_{-i}$ if for every action $b_i \in A_i$ we have that $\pi_i(a_i, a_{-i}) \geqslant \pi_i(b_i, a_{-i})$. An action-tuple (strategy profile) $a^* \in A$ is a *Nash equilibrium* of the game (A, π) if for every player $i \in N$ the action a_i is a best response to a_{-i}, i.e., for every action $b_i \in A_i$ we have $\pi_i(a^*) \geqslant \pi_i(b_i, a^*_{-i})$.

We denote $a_{-i,j} = (a_1, a_2, \ldots, a_{i-1}, a_{i+1}, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n) \in A_{-i,j}$, where

$$A_{-i,j} = \prod_{k \neq i,j} A_k$$

is the actions selected by players other than i and j. Further, $a_{-i,j} = a_{-j,i}$.

A function $Q \colon A \to \mathbb{R}$ is an *exact potential* of the non-cooperative game (A, π) on the player set N if for every player $i \in N$, action tuple $a \in A$ and action $b_i \in A_i$,

$$\pi_i(a) - \pi_i(b_i, a_{-i}) = Q(a) - Q(b_i, a_{-i}). \tag{1}$$

The notion of a game theoretic (exact) potential was seminally introduced in Monderer and Shapley (1996). The potential provides a very strong relationship between the various strategy tuples in the game; it "directs" the game towards its Nash equilibria (in pure strategies). Games that admit exact potentials have very powerful properties, including the guaranteed existence of a Nash equilibrium in pure strategies and the convergence of any improving dynamic improvement process to a Nash equilibrium in a finite number of steps.

A function $Q\colon A \to \mathbb{R}$ is an *ordinal potential* of the non-cooperative game (A, π) on the player set N if for every player $i \in N$, action tuple $a \in A$ and action $b_i \in A_i$,

$$\pi_i(a) - \pi_i(b_i, a_{-i}) > 0 \quad \text{if and only if} \quad Q(a) - Q(b_i, a_{-i}) > 0. \tag{2}$$

(A, π) is an ordinal potential game if it admits an ordinal potential. The notion of an ordinal potential game was also introduced in Monderer and Shapley (1996). Subsequently, Voorneveld and Norde (1997) characterized such games using weak improvement processes. Also, every weighted potential game is an ordinal potential game though the converse need not be true. So any property that holds for ordinal potential games would automatically hold for exact and weighted potential games.

An ordinal (exact) *potential maximizer* is an action tuple $a \in A$ that maximizes the ordinal (exact) potential function Q, i.e., $Q(a) \geqslant Q(b)$ for every $b \in A$. It is obvious that each ordinal potential maximizer is a Nash equilibrium and, hence the notion of a potential maximizer is a refinement of the Nash equilibrium concept.

2.2 Cooperative Games

A *cooperative game* is fully characterized by a function $v\colon 2^N \to \mathbb{R}$ that assigns a productive (output) value $v(S)$ to a coalition $S \subset N$ such that $v(\varnothing) = 0$. The function v is also called a *characteristic function*. An allocation $x \in \mathbb{R}^N$ is *feasible* for the cooperative game v if $\sum_{i \in N} x_i = v(N)$.

The main aim of cooperative game theory is the study of allocation concepts, in particular axiomatic allocation concepts. The main axiomatic allocation concept is the "value", seminally developed and characterized by Shapley (1953). The *Shapley value* is given by the vector $\phi(v) = \{\phi_i(v)\}_{i \in N} \in \mathbb{R}^N$ where

$$\phi_i(v) = \sum_{S \subset N : S \ni i} \frac{(|S|-1)!(n-|S|)!\,[v(S) - v(S \backslash \{i\})]}{n!}. \tag{3}$$

Also, from this it can be shown that

$$\phi_i(v) - \phi_j(v) = \sum_{S \subset N \backslash \{ij\}} \frac{(|S|)!(n-|S|-2)!\,[v(S \cup \{i\}) - v(S \cup \{j\})]}{(n-1)!}. \tag{4}$$

Consider the scenario that players successively join a certain coalition in a given order. One could then assign to each player her marginal contribution to the value of the coalition that is created up till the moment of her entry. The Shapley value of that player can be characterized as the average of her marginal payoffs for all possible orders of entry. (Shapley 1971)

2.3 Networks

In this subsection we define the formal elements to describe networks. Let $N = \{1, 2, \ldots, n\}$ be a finite set of players. Two distinct players $i, j \in N$ with $i \neq j$ are *linked* if i and j are related in some capacity. Usually we think of such links as economically productive relationships between players and can either be formal contracts or informal non-binding agreements. These relationships are *undirected* in the sense that the two players forming a relationship are equals within that relationship. We do not rule out that these relationships have spillover effects on the productive relations between other players. This is captured by the formal description of such network benefits.

Formally, an (undirected) link between i and j is defined as the binary set $\{i, j\}$. Throughout we use the short-hand notation ij to denote the link $\{i, j\}$. It should be clear that ij is completely equivalent to ji.

In total there are $\frac{1}{2}n(n-1)$ possible links on the player set N. The collection of these links on N is denoted by

$$g_N = \{ij \mid i, j \in N \text{ and } i \neq j\}. \tag{5}$$

A *network* g is now defined as an arbitrary collection of links $g \subset g_N$. The collection of all networks on N is denoted by $\mathbb{G}^N = \{g \mid g \subset g_N\}$ and consists of $2^{\frac{1}{2}n(n-1)}$ networks. The network g_N consisting of all possible links is called the *complete network* on N and the network $g_0 = \varnothing$ consisting of no links is denoted as the *empty network*.

For every network $g \in \mathbb{G}^N$ and every player $i \in N$ we denote i's *neighborhood* in g by $N_i(g) = \{j \in N \mid j \neq i \text{ and } ij \in g\}$. Player i therefore is participating in the links in her *link set* $L_i(g) = \{ij \in g \mid j \in N_i(g)\} \subset g$. Let $L_i = L_i(g_N)$ denote the set of all possible links involving player i. We also define $N(g) = \cup_{i \in N} N_i(g)$ and let $n(g) = \# N(g)$ with the convention that if $N(g) = \varnothing$, we let $n(g) = 1$.[2] Also, $n_i(g) = \# N_i(g)$.

Let for all $ij \in g$, $g - ij = g \backslash \{ij\}$. That is, $g - ij$ is the network that remains after removing an existing link ij from g. Similarly, $g + ij = g \cup \{ij\}$ for all $ij \notin g$. Namely, $g + ij$ is the network formed by adding a new link ij to the network g. Furthermore,

[2] We emphasize here that if $N(g) \neq \varnothing$, we have that $n(g) \geqslant 2$. Namely, in those cases the network has to consist of at least one link.

for every set of links $h \subset g$ we denote by $g - h = g \setminus h$ the network resulting after removal of all links in h. Similarly, for $h \subset g_N$ with $h \cap g = \varnothing$, we denote $g + h = g \cup h$ as the network resulting from adding the links in h to g.

Finally, two networks g_1 and g_2 are *adjacent* if they differ by one link, namely, either $g_2 = g_1 - ij$ for some $ij \in g_1$ or $g_1 = g_2 - ij$ for some $ij \in g_2$.

A *path* in g connecting i and j is a set of distinct players $\{i_1, i_2, \dots, i_p\} \subset N(g)$ with $p \geqslant 2$ such that $i_1 = i$, $i_p = j$, and $\{i_1 i_2, i_2 i_3, \dots, i_{p-1} i_p\} \subset g$. Any network for which a path exists between any two players is said to be connected. The network $g' \subset g$ is a *component* of g if for all $i \in N(g')$ and $j \in N(g')$, $i \neq j$, there exists a path in g' connecting i and j and for any $i \in N(g')$ and $j \in N(g)$, $ij \in g$ implies $ij \in g'$. In other words, a component is simply a maximally connected subnetwork of g. We denote the class of network components of the network g by $C(g)$. The set of players that are not connected in the network g are collected in the set of (fully) disconnected players in g denoted by

$$N_0(g) = N \setminus N(g) = \{i \in N \mid N_i(g) = \varnothing\}.$$

Such players are also known as "singletons".

2.4 Link-based stability concepts for networks

We first introduce stability concepts that allow for adding and breaking links separately before considering them together. Note that the stability concepts introduced below are based on the properties of the network itself rather than strategic considerations of the players. This latter viewpoint has been introduced seminally by Jackson and Wolinsky (1996) and is further advocated in Jackson and Watts (2002b), Jackson (2005b), and Bloch and Jackson (2004).

Network formation is based on the net benefits that are generated for the participants in such a network. Formally, we introduce a *network payoff function* as a function $\varphi \colon \mathbb{G}^N \to \mathbb{R}^N$. To every player $i \in N$ the function φ assigns a payoff $\varphi_i(g)$ for participating in the network $g \subset g_N$. This payoff can be positive as well as negative and captures the widespread externalities of the network on the participating players.

Definition 2.1 *Let φ be a network payoff function on the player set N.*

(a) *A network $g \subset g_N$ is **link deletion proof** for φ if for every player $i \in N$ and every $j \in N_i(g)$ it holds that $\varphi_i(g) \geqslant \varphi_i(g - ij)$.*
Denote by $\mathcal{D}(\varphi) \subset \mathbb{G}^N$ the family of link deletion proof networks for φ.

(b) *A network* $g \subset g_N$ *is* ***strong link deletion proof*** *for* φ *if for every player* $i \in N$ *and every* $h \subset L_i(g)$ *it holds that* $\varphi_i(g) \geqslant \varphi_i(g - h)$.
Denote by $\mathcal{D}_s(\varphi) \subset \mathbb{G}^N$ *the family of strong link deletion proof networks for* φ.

(c) *A network* $g \subset g_N$ *is* ***link addition proof*** *if for all players* $i, j \in N$ *with* $ij \notin g$*:* $\varphi_i(g + ij) > \varphi_i(g)$ *implies* $\varphi_j(g + ij) < \varphi_j(g)$.
Denote by $\mathcal{A}(\varphi) \subset \mathbb{G}^N$ *the family of link addition proof networks for* φ.

(d) *A network* $g \in \mathbb{G}^N$ *is* ***strict link addition proof*** *for* $\varphi \colon \mathbb{G}^N \to \mathbb{R}$ *if for all* $i, j \in N\colon\ ij \notin g$ *implies that* $\varphi_i(g + ij) \leqslant \varphi_i(g)$ *as well as* $\varphi_j(g + ij) \leqslant \varphi_j(g)$.
Denote by $\mathcal{A}_s(\varphi) \subset \mathbb{G}^N$ *the family of strict link addition proof networks for* φ.

The two link deletion proofness notions are based on the severance of links in a network by individual players. In particular, the notion of link deletion proofness considers the stability of a network with regard to the deletion of a *single* link. (This concept has been introduced seminally in Jackson and Wolinsky (1996).) Strong link deletion proofness considers the possibility that a player can delete any subset of her existing links. Clearly, strong link deletion proofness implies link deletion proofness. For further details on this concept we refer to Gilles, Chakrabarti, Sarangi, and Badasyan (2005) and Bloch and Jackson (2004).

Similarly, link addition proofness (Jackson and Wolinsky 1996) considers the addition of a single link by two consenting players to an existing network. A network is link addition proof if for every pair of non-linked players, if one of these two players has positive benefits from adding a link between them, the other player only has negative benefits from this addition. Hence, in a network requiring consent this link will never be added.

Strict link addition proofness requires that for every pair of non-linked players, both of these players have non-positive benefits from adding a link between them, i.e., it imposes that neither player has strictly positive incentives to add a link. This formulation does not explicitly introduce a consent requirement; it simply imposes that additional links do not lead to additional revenues. Since both players involved with the formation of an additional link perceive that they individually reduce their payoffs by forming the link, this is a significant strengthening of the link addition proofness requirement. Strict link addition proofness has been introduced in Gilles and Sarangi (2005a) and is further analyzed in Gilles and Sarangi (2005b).

The simplest notion combining both addition and deletion proofness was seminally introduced by Jackson and Wolinsky (1996) through the concept of pairwise stability.

Given that these two conditions can be strengthened in various ways it is also possible to define a variety of modifications of the pairwise stability concept:

Definition 2.2 *Let φ be a network payoff function on the player set N.*

(a) *A network $g \in \mathbb{G}^N$ is **pairwise stable** for φ if g is link deletion proof as well as link addition proof.*
Denote by $\mathcal{P}(\varphi) = \mathcal{D}(\varphi) \cap \mathcal{A}(\varphi) \subset \mathbb{G}^N$ the family of pairwise stable networks for the payoff function φ.

(b) *A network $g \in \mathbb{G}^N$ is **strongly pairwise stable** for φ if g is strong link deletion proof as well as link addition proof.*
Denote by $\mathcal{P}_s(\varphi) = \mathcal{D}_s(\varphi) \cap \mathcal{A}(\varphi) \subset \mathbb{G}^N$ the family of pairwise stable networks for the payoff function φ.

(c) *A network $g \in \mathbb{G}^N$ is **strictly pairwise stable** for φ if g is strong link deletion proof as well as strict link addition proof.*
Denote by $\mathcal{P}^\star(\varphi) = \mathcal{D}_s(\varphi) \cap \mathcal{A}_s(\varphi) \subset \mathbb{G}^N$ the family of strict pairwise stable networks for the payoff function φ.

We refer to Gilles and Sarangi (2005b) for the discussion of equivalence between these three classes of networks. There it is shown that under a convexity property as well as a sign uniformity condition on the network payoff function, all three classes of networks are equal. Gilles and Sarangi (2005a) show in fact that a large sub-family of the class of strictly pairwise stable networks is supported through a learning process based on a myopic belief system. This gives a powerful support to this particular class of networks.

2.5 Myerson's Consent Game

Myerson (1991, page 448) seminally introduced a model of consent in link formation. This model received relatively little serious attention in the literature on network formation until recently. The reason is that the family of Nash equilibria of Myerson's game is very large.[3]

The model introduced by Myerson (1991) is defined as a normal form non-cooperative game and throughout this paper denoted as the *Consent Game*. To define

[3]For a complete characterization of the Nash equilibria in Myerson's consent game we refer to Gilles, Chakrabarti, and Sarangi (2005a). For further discussion we also refer to Jackson (2003) and Jackson (2005b).

this Consent Game, let $\varphi\colon \mathbb{G}^N \to \mathbb{R}^N$ be some network payoff function. Now consider for every player $i \in N$ the strategy set given by

$$A_i = \{(l_{ij})_{j\neq i} \mid l_{ij} \in \{0,1\}\}.$$

Here we interpret $l_{ij} = 1$ to be a signal from player i to player j that she wants to form a link with j. Here, $l_{ij} = 0$ means that player i does not want to form such a link with player j.

A link is formed if both i and j want to form links, namely if $l_{ij} \cdot l_{ji} = 1$. The resulting network is given by

$$g(l) = \{ij \in g_N \mid l_{ij} \cdot l_{ji} = 1\}.$$

We say that the network $g(l)$ is *supported through* the strategy profile l. The game-theoretic payoff function is now given by $\pi_\varphi = (\pi_{\varphi,i})_{i\in N}\colon A \to \mathbb{R}^N$ where $\pi_{\varphi,i}(l) = \varphi_i(g(l))$. We denote the Consent Game corresponding to the network payoff function φ now by $\Gamma_\varphi = (A, \pi_\varphi)$.

Note that every strategy profile supports an unique network, but that a network may be supported through multiple strategy profiles. We denote the set of all strategy profiles supporting a network g by $A_g \equiv \{l \in A \mid g(l) = g\} \subset A$.

Each network $g \subset g_N$ is however supported through a unique *non-superfluous* strategy profile $i_g \in A_g$ satisfying the requirement that for all pairs $i, j \in N$: $l_{ij} = 1$ implies that $l_{ji} = 1$.

3 Exact Network Potentials

First we define exact network potentials and provide a characterization of network payoff functions that admit such potentials. We also relate these potentials to the exact potentials of the Consent Game. We introduce some required concepts and subsequently show the relationship between these. We begin by defining the notion of Equal Bargaining Power,[4] introduced by Myerson (1977) and further developed by Jackson and Wolinsky (1996) and Jackson (2005a).

Definition 3.1 *Let $\varphi\colon \mathbb{G}^N \to \mathbb{R}^N$ be a network payoff function.*

(a) *The network payoff function φ is said to satisfy the* ***equal bargaining power*** *property if for all g and for all $i, j \in N$ with $ij \in g$,*

$$\varphi_i(g) - \varphi_i(g - ij) = \varphi_j(g) - \varphi_j(g - ij).$$

[4]This terminology follows Jackson and Wolinsky (1996).

(b) *The network payoff function* φ ***admits an exact network potential*** *if* φ *satisfies the equal bargaining power property and there exists a function* $\omega\colon \mathbb{G}^N \to \mathbb{R}$ *such that for all* $g \in \mathbb{G}^N$ *and* $ij \in g$,

$$\omega(g) - \omega(g - ij) = \varphi_i(g) - \varphi_i(g - ij) = \varphi_j(g) - \varphi_j(g - ij) \tag{6}$$

The notion of an exact network potential is closely related to that of a game theoretic exact potential. Note, however, that these network potentials are only defined on the class of network payoff functions satisfying Equal Bargaining Power.

Next we define an equivalent condition on network payoff functions. Consider any network payoff function φ satisfying Equal Bargaining Power. Then we define the function $\theta_\varphi\colon \mathbb{G}^N \times g_N \to \mathbb{R}$ by

$$\theta_\varphi(g, ij) = \varphi_i(g) - \varphi_i(g - ij) = \varphi_j(g) - \varphi_j(g - ij)$$

for all $ij \in g$.

Consider any arbitrary network $g \in \mathbb{G}^N$. Let g consist of k links where obviously $k \leqslant \frac{1}{2}n(n-1)$. Next, we label the links in the network g to order them. Let this ordering be represented by

$$g = \left\{\tau_1^\rho, \tau_2^\rho, \dots, \tau_k^\rho\right\}$$

where τ_m^ρ, $m = 1, 2, \dots, k$, is some link $ij \in g$ between two players $i, j \in N$. This introduces an order or permutation $\rho\colon g \leftrightarrows g$ on the network g. There exist a total of $k!$ such orders on g and we denote the set of all these orders by X^g.

Given an order $\rho \in X^g$, we can construct a sequence of adjacent networks that satisfy the following conditions:

$$\begin{aligned} g_0^\rho &= g_0 = \varnothing; \\ g_m^\rho &= g_{m-1}^\rho + \tau_m^\rho \quad \text{for all } m = 1, \dots, k. \end{aligned}$$

Note that from this it immediately follows that $g_k^\rho = g$ for any order $\rho \in X^g$.

On g, the first and the last network of the constructed sequence of adjacent networks are the same. Hence, the given sequence represents the network formation process leading to g starting from the empty network and adding links one at a time according to the order. Then, for any network payoff function φ that satisfies Equal Bargaining Power, define the *Sums function* $S_\varphi : \mathbb{G}^N \times X^g \to \mathbb{R}$ by

$$S_\varphi(g, \rho) = \sum_{m=1}^{k} \theta_\varphi(g_m^\rho, \tau_m^\rho)$$

for all $g \in \mathbb{G}^N$, $\rho \in X^g$. We can now define the required Sums Property as follows.

Definition 3.2 *A network payoff function φ satisfies the* ***Sums Property*** *if φ satisfies Equal Bargaining Power and if for all $g \in \mathbb{G}^N$,*

$$S_\varphi(g, \rho_1) = S_\varphi(g, \rho_2) \tag{7}$$

for any two arbitrary orders $\rho_1, \rho_2 \in X^g$.

Our main equivalence theorem for exact network potentials is stated below and provides a complete characterization of the existence of exact network potentials in terms of the Consent Game as well as the Sums Property.

Theorem 3.3 *The following statements are equivalent for an arbitrary network payoff function φ on $\mathbb{G}^N$.*

(i) *The network payoff function φ admits an exact network potential.*

(ii) *The Consent Game Γ_φ admits an exact potential.*

(iii) *The network payoff function φ satisfies the Sums Property.*

The proof of Theorem 3.3 is relegated to the appendix.

The following example illustrates a familiar case that satisfies the properties of Theorem 3.3. Throughout we use the terminology from Gilles, Chakrabarti, and Sarangi (2005b).

Example 3.4 Link-based network payoff functions refer to linear network payoff functions that are devoid of network externalities. Benefits accrue only from direct links and each link yields a fixed benefit or loss independent of network structure. This has been explored by Baron, Durieu, Haller, and Solal (2006) in a slightly more general context. For further details we also refer to Gilles, Chakrabarti, and Sarangi (2005b).

Let $\xi_i \colon L_i \to \mathbb{R}$ be the link benefit function for player $i \in N$ which assigns to each potential link $ij \in L_i$ a payoff $\xi_i(ij) \in \mathbb{R}$. Based on this link-based payoff function one can define a network payoff function φ_ξ where $\varphi_{\xi,i}(g) = \sum_{j \in N_i(g)} \xi_i(ij)$. If for all $ij \in g_N$, it holds that $\xi_i(ij) = \xi_j(ij) = \xi(ij)$, then the network payoff function is referred to as a *mutual* link-based network payoff function.

Such mutual link-based payoff functions admit an exact network potential. An exact network potential for such a mutual link-based payoff function φ_ξ is in fact given by

$$\omega_\xi(g) = \sum_{ij \in g} \xi(ij).$$

The case that φ_ξ is not a mutual link-based network payoff function does not admit an exact network potential. For further discussion we refer to Example 4.2. □

Next we elaborate on the relationship between network payoff functions admitting exact network potentials and the Shapley value of a constructed cooperative game. Similar relationships have been explored by Monderer and Shapley (1996), Qin (1996), and Ui (2000). Given $\varphi\colon \mathbb{G}^N \to \mathbb{R}^N$, we introduce for every network $g \in \mathbb{G}^N$ a cooperative game $U_{\varphi,g}\colon 2^N \to \mathbb{R}$, where for all $S \subset N$,

$$U_{\varphi,g}(S) = \sum_{i \in S} \varphi_i(g|_S)$$

where $g|_S = \{ij \in g \mid i \in S \text{ and } j \in S\}$. In other words, the characteristic function is such that the value generated by a coalition is the sum of the payoffs of all members of the coalition for the sub-network in which all links for which one or both members pertaining to any link are outside the coalition is removed. The Shapley value of this cooperative game is denoted by $\phi(U_{\varphi,g}) = \{\phi_i(U_{\varphi,g})\}_{i \in N}$.

Definition 3.5 *The network payoff function φ is **Shapley-consistent** if for every $g \in \mathbb{G}^N$ it holds that $\varphi(g) = \phi(U_{\varphi,g})$.*

In other words, if payoffs were redistributed according to the Shapley value, the payoffs would not change. Subsequently we introduce a class of network payoff functions where if a player is an unconnected singleton, she earns the same payoff irrespective of network structure.

Definition 3.6 *We define $\widetilde{\Phi}$ to be the family of network payoff functions $\varphi\colon \mathbb{G}^N \to \mathbb{R}^N$ such that for every player $i \in N$, there exists some $\tau_i \in \mathbb{R}$ such that for every network $g \in \mathbb{G}^N$ with $i \in N_0(g)$ it holds that $\varphi_i(g) = \tau_i$.*

The class $\widetilde{\Phi}$ includes all network payoff functions that do not display widespread externalities beyond the connected players $N(g) \subset N$. Namely, every player $i \in N_0(g) = N \setminus N(g)$ receives exactly the same payoff irrespective of the structure of the connected part of the network g.

The next result extends the insights presented in Qin (1996) and Ui (2000) for the setting of networks with widespread externalities. For a proof of this theorem we again refer to the appendix.

Theorem 3.7 *The two following statements hold for any network payoff function φ.*

(a) *If φ is Shapley-consistent, then φ admits an exact network potential.*

(b) *If $\varphi \in \widetilde{\Phi}$ admits an exact network potential, then φ is Shapley-consistent.*

We remark that Theorem 3.7 is similar to the main result of Qin (1996). Qin proved his result for so-called *cooperation structures* defined as cooperative games in which only connected groups of players within a communication network can form coalitions.[5] Our result extends Qin's insights to the setting of social networks in the presence of widespread externalities to communication and cooperation.

If $\varphi \notin \widetilde{\Phi}$, then the converse of Theorem 3.7(a) does not hold. Below, we discuss a simple counter-example.

Example 3.8 Let $n = 3$ and consider the network payoff function given in the following table:

Network	$\varphi_1(g)$	$\varphi_2(g)$	$\varphi_3(g)$
$g_0 = \varnothing$	0	0	0
$g_1 = \{12\}$	1	1	2
$g_2 = \{13\}$	1	0	1
$g_3 = \{23\}$	0	1	1
$g_4 = \{12, 13\}$	3	2	4
$g_5 = \{12, 23\}$	2	3	4
$g_6 = \{13, 23\}$	2	2	3
$g_7 = g_N$	5	5	7

From the table one can deduce that $\varphi \notin \widetilde{\Phi}$. Indeed, player 3 receives $\varphi_3(g_1) = 2 \neq 0 = \varphi_3(g_0)$ even though $3 \in N_0(g_1)$.

The reader can verify that the Sums Property is satisfied. For every network g we may compute the Sums function $S_\varphi(g)$, which is independent of the selected order on g, and remark that S_φ is in fact an exact network potential for this network payoff function. (This is shown formally in the proof of Theorem 3.3.)

$$
\begin{aligned}
S_\varphi(g_0) &= 0 \\
S_\varphi(\{12\}) &= S_\varphi(\{13\}) = S_\varphi(\{23\}) = 1 \\
S_\varphi(\{12, 13\}) &= S_\varphi(\{12, 23\}) = S_\varphi(\{13, 23\}) = 3 \\
S_\varphi(\{12, 13, 23\}) &= 6
\end{aligned}
$$

In fact a reinterpretation of this particular network payoff function φ is that it is a simple construction direct from the network potential S_φ except for the anomalous value of $\varphi_3(g_1)$.

[5]The notion of such a communication requirement on coalition formation was seminally introduced by Myerson (1977). We remark here that in this approach the network does not generate externalities, but acts as a limiting device on the cooperative possibilities.

Now one can verify that for the network g_1 the constructed cooperative game $U = U_{\varphi,g_1}$ is given by

$$
\begin{aligned}
&U(\varnothing) = U(1) = U(2) = U(3) = 0\\
&U(12) = 2\\
&U(13) = U(23) = 0\\
&U(123) = 4
\end{aligned}
$$

This cooperative game has a Shapley value given by $\phi(U) = (1\frac{2}{3}, 1\frac{2}{3}, \frac{2}{3}) \neq (1, 1, 2) = \varphi(g_1)$. Thus, this payoff function—although admitting an exact network potential—does not satisfy Shapley consistency. □

4 Ordinal Network Potentials

We proceed to discuss ordinal network potentials and explore their relationship with ordinal potentials of the corresponding Consent Game. Ordinal network potentials generalize the notion of an exact network potential from cardinal structures to ordinal structures similarly as seminally introduced by Monderer and Shapley (1996). We subsequently examine the properties of network payoff functions admitting such ordinal network potentials.

Like we did for the case of exact network potentials, we are only able to introduce the notion of an ordinal network potential for a class of network payoff functions that satisfies the so-called pairwise sign compatibility condition, which generalizes the equal bargaining power requirement.

Definition 4.1 *Let $\varphi\colon \mathbb{G}^N \to \mathbb{R}^N$ be a network payoff function.*

(a) *φ is said to satisfy **pairwise sign compatibility** (PSC) if for all g and $ij \in g$, the following three properties hold:*

$$
\begin{aligned}
\varphi_i(g) > \varphi_i(g - ij) &\implies \varphi_j(g) > \varphi_j(g - ij);\\
\varphi_i(g) < \varphi_i(g - ij) &\implies \varphi_j(g) < \varphi_j(g - ij);\\
\varphi_i(g) = \varphi_i(g - ij) &\implies \varphi_j(g) = \varphi_j(g - ij).
\end{aligned}
$$

(b) *φ **admits an ordinal network potential** if φ satisfies pairwise sign compatibility and there exists a function $\omega\colon \mathbb{G}^N \to \mathbb{R}$ such that for all g and $ij \in g$,*

$$\omega(g) > \omega(g - ij) \iff \varphi_i(g) > \varphi_i(g - ij); \tag{8}$$

$$\omega(g) < \omega(g - ij) \iff \varphi_i(g) < \varphi_i(g - ij); \tag{9}$$

$$\omega(g) = \omega(g - ij) \iff \varphi_i(g) = \varphi_i(g - ij). \tag{10}$$

We introduce a short-hand notation for denoting the above, namely, (8), (9) and (10) are equivalent to $\text{Sign}\,(\omega(g), \omega(g - ij)) = \text{Sign}\,(\varphi_i(g), \varphi_i(g - ij))$.

To illustrate the introduction of the concept of an ordinal network potential we return to the case of link-based network payoffs already discussed in Example 3.4. The following illustrates when ordinal network potentials can be admitted for that particular case.

Example 4.2 Let N be van arbitrary player set. As in Example 3.4 consider for each player $i \in N$ a link-based payoff function $\xi_i \colon L_i \to \mathbb{R}$. Again we define the network payoff function by $\varphi_{\xi,i}(g) = \sum_{ij \in L_i(g)} \xi_i(ij)$.
If for all $i, j \in N$ it holds that $\xi_i(ij) > 0$ implies $\xi_j(ij) > 0$ as well as $\xi_i(ij) < 0$ implies $\xi_j(ij) < 0$, then φ_ξ satisfies PSC and even admits an ordinal network potential. In that case an ordinal potential function is given by

$$\omega_\xi(g) = \sum_{i \in N} \varphi_{\xi,i}(g) = \sum_{ij \in g} [\xi_i(ij) + \xi_j(ij)]$$

It is easy to see that this ordinal network potential generalizes the formulated exact network potential function introduced in Example 3.4. □

The following theorem is the analogue of Theorem 3.3 for the ordinal case. However, only a partial relationship can be developed for the ordinal case. A proof of this theorem can be found in the appendix.

Theorem 4.3 *Let φ be some network payoff function. If the corresponding Consent Game Γ_φ admits an ordinal potential, then φ admits an ordinal network potential.*

The existence of an ordinal network potential is a strictly weaker property than the property that the corresponding Consent Game admits an ordinal potential. The following example shows this. The network payoff function below admits an ordinal network potential but the corresponding Consent Game does not admit an ordinal potential. In fact, it violates the *finite improvement property* of Monderer and Shapley (1996) which is a necessary condition for a non-cooperative game to admit an ordinal potential.[6]

Example 4.4 Let $n = 3$ and consider the following network payoff function.

[6] We omit a discussion of the finite improvement property since it involves concepts that are unrelated to rest of the paper. For details we refer to Monderer and Shapley (1996).

Network	$\varphi_1(g)$	$\varphi_2(g)$	$\varphi_3(g)$
$g_0 = \varnothing$	0	0	0
$g_1 = \{12\}$	-5	-5	-2
$g_2 = \{13\}$	-2	0	-2
$g_3 = \{23\}$	-1	-1	-1
$g_4 = \{12, 13\}$	1	-4	-1
$g_5 = \{12, 23\}$	-2	-3	0
$g_6 = \{13, 23\}$	1	1	0
$g_7 = g_N$	5	5	5

This network payoff function φ admits the following ordinal network potential:

$$\begin{aligned} \omega(g_0) &= 3 \\ \omega(\{12\}) &= 0 \\ \omega(\{13\}) &= 0 \\ \omega(\{23\}) &= 2 \\ \omega(\{12, 23\}) &= 1 \\ \omega(\{12, 13\}) &= 1 \\ \omega(\{13, 23\}) &= 3 \\ \omega(\{12, 13, 23\}) &= 4 \end{aligned}$$

It can also be derived for the given network payoff function φ that the finite improvement property (Monderer and Shapley 1996) is not satisfied for the corresponding Consent Game Γ_φ. Indeed, we device the following four strategy tuples[7] in A:

$$\begin{aligned} &l_1 = (00, 11, 10) \text{ with } g(l_1) = g_0 = \varnothing \\ &l_2 = (11, 11, 10) \text{ with } g(l_2) = g_4 = \{12, 13\} \\ &l_3 = (11, 11, 01) \text{ with } g(l_3) = g_5 = \{12, 23\} \\ &l_4 = (00, 11, 01) \text{ with } g(l_4) = g_3 = \{23\} \end{aligned}$$

We claim that $l_1 - l_2 - l_3 - l_4 - l_1$ is an improvement cycle in the sense of Monderer and Shapley (1996). Indeed, Player 1 increases his individual payoff π_1 by moving from l_1 (g_0) to l_2 (g_4) and from l_3 (g_5) to l_4 (g_3). Also, Player 3 increases his individual payoff π_3 by moving from l_2 (g_4) to l_3 (g_5) and from l_4 (g_3) to l_1 (g_0). This shows that indeed the network payoff function φ generates a Consent Game Γ_φ that does not satisfy the finite improvement property and, therefore, does not admit an

[7] In this example we use shorthand notation for these strategy tuples. We use the listing convention given by $l = (l_{12}\, l_{13}, l_{21}\, l_{23}, l_{31}\, l_{32})$.

ordinal potential.

The underlying motive is mainly that non-cooperative game theoretic improvement paths are very different in nature from link-based improvement paths as introduced in the next section. Indeed, individual players are completely in control of the strategies that they select. This is fundamentally different from the coordinated strategic modification underlying a link-based improvement. □

In the next section we discuss several properties of network payoff function which admit ordinal network potentials and those for which the corresponding Consent Game admits an ordinal potential. We introduce the following three classes of network payoff functions:

Φ_1: This is the set of network payoff functions φ that admit an exact network potential. This is equivalent to the requirement that the corresponding Consent Game Γ_φ admits an exact potential.

Φ_2: This is the set of network payoff functions φ for which the corresponding Consent Game Γ_φ admits an ordinal potential.

Φ_3: This is the set of network payoff functions φ that admit an ordinal network potential.

From the above definitions, it can be concluded that

$$\varnothing \neq \Phi_1 \subsetneq \Phi_2 \subsetneq \Phi_3.$$

We emphasize that in principle there is no link of the sets introduced above with the collection $\widetilde{\Phi}$ introduced in the preceding section of this paper.

5 Potentials and network stability

In this section we discuss the relationship between the existence of an ordinal network potential and the convergence of improvement paths to a pairwise stable network in the sense of Jackson and Watts (2002a). Subsequently we consider similar questions with regard to network payoff functions in the class Φ_2 for which the corresponding Consent Game admits an ordinal potential. We show that these payoff functions imply stronger convergence and existence properties. (Jackson and Watts 2001)

5.1 Improvement processes

We first introduce some concepts taken from Jackson and Watts (2002a) and Jackson and Watts (2002b).[8]

Definition 5.1 *An* ***improvement path*** *from a network* g *to a network* g' *is a finite sequence of adjacent networks* $g_1, g_2, \dots, g_K$ *with* $g_1 = g$ *and* $g_K = g'$ *such that for any* $k \in \{1, \dots, K-1\}$ *either*

(i) $g_{k+1} = g_k - ij$ *for some* ij *such that* $\varphi_i(g_k - ij) > \varphi_i(g_k)$*, or*

(ii) $g_{k+1} = g_k + ij$ *for some* ij *such that* $\varphi_i(g_k + ij) > \varphi_i(g_k)$ *and* $\varphi_j(g_k + ij) \geqslant \varphi_j(g_k)$.

An improvement path $C \subset \mathbb{G}^N$ *is an* ***improvement cycle*** *if for any* $g \in C$ *and* $g' \in C$ *there exists an improvement path from* g *to* g'.
If there is an improvement path from g *to* g' *and* g *and* g' *are adjacent, then we also say that* g' ***defeats*** g*. It should be clear that every pairwise stable network is undefeated.*

It is obvious that if there are no improvement cycles, then the network payoff function admits a pairwise stable network. This is because in general there are three possibilities. First, there are no improvement paths, in which case every network is pairwise stable. Second, every improvement path terminates in some pairwise stable network. Third, there is at least one improvement path that does not terminate. Given that there are only a finite number of networks, the latter simply means there exists an improvement cycle. Therefore, the absence of improvement cycles guarantees existence of at least one pairwise stable network. We formalize this using the following insight, which forms the foundation of the research on pairwise stable networks of Jackson and Watts (2002a) and Jackson and Watts (2001).

Lemma 5.2 : Corollary to Jackson and Watts (2001, Lemma 1).
Given any network payoff function, there either exists at least one improvement cycle or at least one pairwise stable network.

In fact, Jackson and Watts (2001) prove a much stronger result, namely that there exists at least one closed improvement cycle or at least one pairwise stable network, a closed improvement cycle being an improvement cycle such that there are no alternative improvement paths emanating from any of the networks that are part of

[8]The terminology of an "improvement path" introduced here should be distinguished strictly from the concept with the same name in non-cooperative game theory. The notion developed here refers to a sequence of networks with certain properties rather than a sequence of strategy tuples as used in non-cooperative game theory.

the improvement cycle. Jackson and Watts (2001) have also defined necessary and sufficient conditions for absence of improvement cycles.

Next we turn to the discussion of the relationship between improvement paths and the presence of ordinal network potentials.

Theorem 5.3 *If $\varphi \in \Phi_3$ admits an ordinal network potential, then the following properties hold:*

(a) *There exists at least one pairwise stable network;*

(b) *There are no improvement cycles;*

(c) *The set of strongly pairwise stable and strictly pairwise stable networks coincide.*

For a proof of Theorem 5.3 we refer to the appendix.

The converse of assertion 5.3(b) is not true. There can be payoff functions without improvement cycles, but for which there does not exist an ordinal network potential. The next counter-example is a modification of an example developed in Jackson and Watts (2001).

Example 5.4 Let $n = 3$. Suppose the network payoff function is such that $\{12, 23, 13\}$ defeats $\{12, 23\}$ defeats $\{12\}$ defeats $\{12, 13\}$, but that players 2 and 3 are both indifferent between $\{12, 23, 13\}$ and $\{12, 13\}$. Suppose also no other network defeats any other. Here the reader can verify that there are no improvement cycles. But we can show that no ordinal network potential exists.

Suppose by contradiction that an ordinal network potential, say ω, exists. Then, by definition since $\{12, 23, 13\}$ defeats $\{12, 23\}$ defeats $\{12\}$ defeats $\{12, 13\}$, it has to hold that

$$\begin{aligned} \omega(\{12, 23, 13\}) &> \omega(\{12, 23\}) \\ \omega(\{12, 23\}) &> \omega(\{12\}) \\ \omega(\{12\}) &> \omega(\{12, 13\}) \end{aligned}$$

Hence, $\omega(\{12, 23, 13\}) > \omega(\{12, 13\})$. But the fact that 2 and 3 are both indifferent between $\{12, 23, 13\}$ and $\{12, 13\}$ implies

$$\omega(\{12, 23, 13\}) = \omega(\{12, 13\})$$

which is a contradiction. $\square$

Next, we show that two additional conditions, Pairwise Sign Compatibility and a "no indifference" condition guarantee the existence of ordinal network potential. In the above case, no indifference is violated. The used terminology follows Jackson and Watts (2001).

The network payoff function *exhibits no indifference* if for any two adjacent networks g and g', either g defeats g' or g' defeats g.

We can now establish the sufficient conditions for existence of ordinal network potentials. For a proof of the next theorem we refer as usual to the appendix.

Theorem 5.5 *If the network payoff function satisfies Pairwise Sign Compatibility, exhibits no indifference, and is such that there are no improvement cycles, then the network payoff function admits an ordinal network potential and hence belongs to* Φ_3*.*

5.2 Existence of strictly pairwise stable networks

Next we address the implications of the existence of potentials for the existence of pairwise stable networks. As Jackson and Watts (2001) show, improvement paths in general converge to pairwise stable networks. Hence, if convergence can be established, the existence of such networks is guaranteed as well. This is closely related to the presence of potential functions as originally pointed out by Jackson and Watts.

Here we generalize the findings of Jackson and Watts. In particular, we are able to show that for the subclass Φ_2 of network payoff functions for which the corresponding Consent Game has an ordinal potential, we can guarantee the existence of *strictly* pairwise stable networks. This rather strong existence result implies, therefore, the existence of strongly pairwise stable as well as regular pairwise stable networks.

Theorem 5.6 *If* $\varphi \in \Phi_2$*, then there exists at least one strictly pairwise stable network.*

The proof of Theorem 5.6 is based on the fact that, given that there are a finite number of strategy tuples in the corresponding Consent Game, for any $\varphi \in \Phi_2$ there exists an ordinal potential maximizer. This potential maximizer corresponds to a strictly pairwise stable network.

This raises the question whether every strictly or strongly pairwise stable network is supported through a potential maximizer of the Consent game. The next counterexample shows this is not the case.

Example 5.7 Let $n = 3$ and consider the network payoff function be given by

$\varphi_k(g) = 1$ for all $k \in N$ if $g = \{ij, ih\}$ has a line topology, and

$\varphi_i(g) = 0$ for all $i \in N$ and for all other $g \in \mathbb{G}^N$.

This network payoff function admits an exact network potential, i.e., $\varphi \in \Phi_1$ and, hence, $\varphi \in \Phi_2$.[9] Now consider the empty network g_0. It is strongly pairwise stable, and, consequently, strictly pairwise stable as well. Indeed, there are no links to be deleted and addition of any one link will not change payoffs.
Next, we show that a strategy tuple supporting g_0 is not necessarily a potential maximizer. First, consider the strategy tuple l given by $l = (00, 10, 10)$. It supports the empty network. But it is not an ordinal potential maximizer. Let Q be an arbitrary ordinal potential of the corresponding Consent Game (A, π). Then, define $\widehat{l}_1 \in A_1$ to be such that $\widehat{l}_{12} = \widehat{l}_{13} = 1$, then

$$\pi_1(\widehat{l}_1, l_{-1}) = 1 > 0 = \pi_1(l) \Rightarrow Q(\widehat{l}_1, l_{-1}) > Q(l)$$

and hence l is not an ordinal potential maximizer. In fact, we can extend this reasoning and show that none of the strategy tuples supporting the empty network g_0 are ordinal potential maximizers. $\square$

6 Some applications

In this section we discuss several applications of the theory of network potentials developed in the previous sections. First we turn shortly to a well-known application from social network theory, the so-called *connections model*, due to Jackson and Wolinsky (1996).

Subsequently we turn to the discussion of the introduction of costs in the link formation process. (Gilles, Chakrabarti, and Sarangi 2005a) If link formation is costly, we show that essentially the main implication of the presence of network potentials is not affected. Under both two-sided and one-sided link formation costs, the presence of a network potential implies that the Consent Game admits a potential as well. It should be clear that the reverse of this implication can no longer be guaranteed as is the case for costless link formation.

6.1 The connections model

We illustrate Theorem 4.3 with an application that illustrates the applicability of ordinal network potentials to models considered in the literature. The connections model describes that communication over multiple links reduces the resulting benefits exponentially. Besides social applications, the connections model has application in

[9] In fact, an exact network potential function ω for φ is given by $\omega(g) = 1$ for $g = \{ij, ik\}$ with a line topology and $\omega(g) = 0$ for all other $g \in \mathbb{G}^N$.

engineering, in particular wireless mobile ad-hoc networks. (Srivastava, Neel, Hicks, MacKenzie, Lau, DaSilva, Reed, and Gilles 2005)

Let the player set N be arbitrary and consider any $g \in \mathbb{G}^N$. The connections network payoff function for player i in network g is now given by

$$\varphi_i^\delta(g) = \sum_{j \neq i} \delta^{t_{ij}(g)} - \sum_{j:\ ij \in g} c_{ij} \tag{11}$$

where $0 < \delta < 1$, $c_{ij} \geqslant 0$ is the cost of establishing link ij for player i, and $t_{ij}(g)$ is the number of links on the shortest path between i and j. The number $t_{ij}(g)$ is also called the "geodesic distance" between i and j in network g. (If i and j are not linked, i.e., $ij \notin g$, then $t_{ij}(g) = \infty$ by convention.) If for the connections model given in (11) it holds that all link formation costs are equal, i.e., $c_{ij} = c \geqslant 0$, then we refer to this setup as the *symmetric* connections model.

In the symmetric connections model, if $c < \delta - \delta^2$, then φ^δ admits an ordinal network potential. In that case, any link formed increases the payoffs of both players forming the link and does not reduce the payoffs of all the other players. Hence, by definition, $\omega(g) = \sum_{i \in N} \varphi_i^\delta(g)$ is an ordinal network potential.

6.2 Two-sided link formation costs

Let $N = \{1, \ldots, n\}$ be a given set of players and $\psi \colon \mathbb{G}^N \to \mathbb{R}_+^N$ be a fixed, but arbitrary network benefit function representing the gross benefits that accrue to the players in a network.[10] For every player $i \in N$ we introduce individualized link formation costs represented by $c_i = (c_{ij})_{j \neq i} \in \mathbb{R}_+^{N \setminus \{i\}}$. (Note that for some links $ij \in g_N$ it might hold that $c_{ij} \neq c_{ji}$.) Thus, the pair $\langle \psi, c \rangle$ represents the basic payoffs and costs of network formation to the individuals in N.

As in the standard Consent Game we let for every player $i \in N$ her action set be given by

$$A_i = \{(l_{ij})_{j \neq i} \mid l_{ij} \in \{0, 1\}\}$$

As before, player i seeks contact with player j if $l_{ij} = 1$ and a link is formed if both players seek contact, i.e., $l_{ij} = l_{ji} = 1$.

Let $A = \prod_{i \in N} A_i$ where $l \in A$. Then the resulting network is given by

$$g(l) = \{ij \in g_N \mid l_{ij} = l_{ji} = 1\}.$$

[10]We emphasize that a network benefit function is simply a network payoff function that only admits non-negative payoffs to the players.

Now, however, link formation is costly. Approaching player j to form a link costs player i an amount $c_{ij} \geqslant 0$. This results in the following game theoretic payoff function for player i:

$$\pi_i^a(l) = \psi_i(g^a(l)) - \sum_{j \neq i} l_{ij} \cdot c_{ij} \tag{12}$$

where c is the link formation cost introduced above.

The pair $\langle \psi, c \rangle$ thus generates the non-cooperative game $\Gamma_2(\psi, c) = (A, \pi^a)$ as described above. The game $\Gamma_2(\psi, c)$ is denoted as the *Consent Game under two-sided link formation costs*.

Proposition 6.1 *Consider the Consent Game under two-sided link formation costs $\Gamma_2(\psi, c)$ based on $\langle \psi, c \rangle$. If $\psi \in \Phi_1$ is a network benefit function that admits an exact network potential, then the Consent Game under two-sided link formation costs admits an exact potential.*

Proof. Let ψ admit an exact potential $\omega \colon \mathbb{G}^N \to \mathbb{R}$. Now consider the function $Q^a \colon A \to \mathbb{R}$ given by

$$Q^a(l) = \omega(g(l)) - \sum_{ij \in g_N} [\, l_{ij} \cdot c_{ij} + l_{ji} \cdot c_{ji} \,]$$

Using similar arguments as used in the proof of theorem 3.3, we now can show that the function Q^a is in fact an exact potential for the Consent Game under two-sided link formation costs (A, π^a). ■

The assertion of Proposition 6.1 is not true for ordinal network potentials. Indeed consider the following situation. Let there be three players ($n = 3$) and the network benefits function ψ given in the matrix below. We remark that the network benefits function ψ admits an ordinal network potential given by ω, also provided in this matrix.

Network	$\psi_1(g)$	$\psi_2(g)$	$\psi_3(g)$	$\omega(g)$
$g_0 = \varnothing$	0	0	0	1
$g_1 = \{12\}$	5	6	7	3
$g_2 = \{13\}$	8	9	10	5
$g_3 = \{23\}$	11	12	13	7
$g_4 = \{12, 13\}$	13.5	14	14.5	9
$g_5 = \{12, 23\}$	15	15.5	16	11
$g_6 = \{13, 23\}$	16.5	17	18	13
$g_7 = g_N$	20	29	30	15

Next consider costly network formation such that $c_{12} = 4$, $c_{31} = 6$, $c_{23} = 13$, and $c_{ij} = 0$ for all other ordered pairs $ij \notin \{12, 31, 23\}$.

Within the provided setting we can now construct an improvement cycle in the consent game under two-sided link formation costs $\Gamma_2(\psi, c)$, thus showing that the impossibility that this game admits an ordinal potential. This improvement cycle is as follows:

$$l_1 = (11, 11, 11) \text{ with } g(l_1) = g_7 = g_N$$
$$l_2 = (01, 11, 11) \text{ with } g(l_2) = g_5$$
$$l_3 = (01, 11, 01) \text{ with } g(l_3) = g_3$$
$$l_4 = (01, 10, 01) \text{ with } g(l_4) = g_0 = \varnothing$$
$$l_5 = (11, 10, 01) \text{ with } g(l_5) = g_1$$
$$l_6 = (11, 10, 11) \text{ with } g(l_6) = g_4$$

It is easy to see that this strategy sequence indeed constitutes an improvement cycle in $\Gamma_2(\psi, c)$.

This shows the claim that even though the network benefit function ψ admits an ordinal potential, the resulting consent game under two-sided link formation costs does not necessarily have to admit an ordinal potential.

6.3 One-sided link formation costs

In this section we discuss the Consent Game under one-sided link formation costs. Here links are formed by mutual agreement, but only one player in the pair under consideration initiates the link formation process and the other player only responds to this link formation attempt. The initiator incurs the formation costs of the link, while the respondent incurs no costs. Our discussion follows the setup developed in Gilles, Chakrabarti, and Sarangi (2005a).[11] Hence, a different strategy space reflecting the difference between initiator and respondent is called for.

Formally, consider an arbitrary player set N and a non-negative network benefit function $\psi\colon \mathbb{G}^N \to \mathbb{R}^N_+$. To model the separation of the initiation of links and the responding to link formation initiations, we introduce for every player $i \in N$ a strategy set given by

$$A^b_i = \{(l_{ij}, r_{ij})_{j \neq i} \mid l_{ij}, r_{ij} \in \{0, 1\}\}. \tag{13}$$

[11] We remark that a similar link formation structure has been already discussed by Slikker (2000) and Slikker, Gilles, Norde, and Tijs (2005) in the context of the formation of *directed* networks. See also Dutta and Jackson (2000).

Player i acts as the initiator in forming a link with player j if $l_{ij} = 1$. Player j responds positively to this initiative if $r_{ji} = 1$. A link is established if formation is initiated and accepted, i.e., if $l_{ij} = r_{ji} = 1$. This is formalized as follows.

Let $A^b = \prod_{i \in N} A_i^b$ be the corresponding strategy tuple space. Given the link formation procedure described, for any $(l, r) \in A^b$, the resulting network is now given by

$$g^b(l, r) = \{ij \in g_N \mid l_{ij} = r_{ji} = 1\}. \tag{14}$$

When player i initiates the formation of a link with player j she incurs a cost of $c_{ij} \geqslant 0$. Responding to the initiative by another player, however, is costless. This results in the following game theoretic payoff function for player i:

$$\pi_i^b(l, r) = \varphi_i(g^b(l, r)) - \sum_{j \neq i} l_{ij} \cdot c_{ij} \tag{15}$$

where c denotes the link formation costs.

Analogous to the previous model with two-sided link formation costs, the pair $\langle \psi, c \rangle$ now generates the non-cooperative game $\Gamma_1(\psi, c) = (A^b, \pi^b)$ introduced above. The game $\Gamma_1(\psi, c)$ represents the *Consent Game under one-sided link formation costs*.

The next proposition shows only a partial statement of the similar proposition formulated for the Consent Game under two-sided link formation costs. The structure of one-sided link formation costs is based on a fundamental asymmetry that can no longer be linked directly to standard Consent Game Γ_ψ based on the benefit function ψ admitting an ordinal potential.

Proposition 6.2 *Let $\Gamma_1(\psi, c)$ be the Consent Game under one-sided link formation costs based on $\langle \psi, c \rangle$. If $\psi \in \Phi_1$ admits an exact network potential, then the Consent Game under one-sided link formation costs $\Gamma_1(\psi, c)$ admits an exact potential.*

Proof. Let $\psi \in \Phi_1$ admit an exact potential $\omega \colon \mathbb{G}^N \to \mathbb{R}$. Now consider the function $Q^b \colon A^b \to \mathbb{R}$ given by

$$Q^b(l, r) = \omega(g^b(l, r)) - \sum_{ij \in g_N} [\, l_{ij} \cdot c_{ij} + l_{ji} \cdot c_{ji} \,]$$

Again using similar techniques as used in the proof of Theorem 3.3 we can now show that the function Q^b is in fact an exact potential for the Consent Game under one-sided link formation costs $\Gamma_1(\psi, c)$. ■

7 Coda

While potentials are defined for non-cooperative games, a similar notion called network potentials can be defined for network payoff functions. It is closely related to the potentials of a non-cooperative link formation game due to Myerson (1991). An exact network potential exists if and only if the Consent Game admits an exact potential. The existence of network ordinal potentials is however weaker than the existence of ordinal potentials for the Consent Game. When these potentials exist, the network payoff function has attractive properties. When exact network potentials exist, there exists both a strongly pairwise stable and a strictly pairwise stable network. Under the much weaker condition of the existence of ordinal potentials, there are no network improvement cycles, at least one pairwise stable network exists and the set of strongly pairwise stable and strictly pairwise stable networks coincide.

Topics for further research include when does existence of ordinal network potentials guarantee that the Consent Game has an ordinal potential. Another topic of interest is what are the interesting examples of network payoff functions satisfy the somewhat demanding conditions of the existence of exact and ordinal network payoff functions. Gilles, Chakrabarti, and Sarangi (2005b) discuss an example of network payoff function for which an exact network potential exists.

References

BALA, V., AND S. GOYAL (2000): "A Non-Cooperative Model of Network Formation," *Econometrica*, 68, 1181–1230.

BARON, R., J. DURIEU, H. HALLER, AND P. SOLAL (2006): "Complexity and Stochastic Evolution of Dyadic Networks," *Computers & Operations Research*, 33, 312–327.

BLOCH, F., AND M. O. JACKSON (2004): "The Formation of Networks with Transfers among Players," Working Paper, GREQAM, Université d'Aix-Marseille, Marseille, France.

DUTTA, B., AND M. O. JACKSON (2000): "The Stability and Efficiency of Directed Communication Networks," *Review of Economic Design*, 5, 251–272.

DUTTA, B., A. VAN DEN NOUWELAND, AND S. TIJS (1998): "Link Formation in Cooperative Situations," *International Journal of Game Theory*, 27, 245–256.

GILLES, R. P., S. CHAKRABARTI, AND S. SARANGI (2005a): "Social Network Formation with Consent: Nash Equilibria and Pairwise Refinements," Working paper, Department of Economics, Virginia Tech, Blacksburg, VA.

——— (2005b): "Stability and Link-Based Network Payoffs," Working paper, Department of Economics, Virginia Tech, Blacksburg, VA.

GILLES, R. P., S. CHAKRABARTI, S. SARANGI, AND N. BADASYAN (2005): "Network Intermediaries," Working Paper, Department of Economics, Virginia Tech, Blacksburg, Virginia.

GILLES, R. P., AND S. SARANGI (2005a): "Building Social Networks," Working Paper, Department of Economics, Virginia Tech, Blacksburg, VA.

——— (2005b): "Stable Networks and Convex Payoffs," CentER Discussion Paper 2005-84, Tilburg University, Tilburg, the Netherlands.

HART, S., AND A. MAS-COLELL (1989): "Potential, Value, and Consistency," *Econometrica*, 57, 589–614.

JACKSON, M. O. (2003): "the Stability and efficiency of economic and Social Networks," in *Networks and Groups: Models of Strategic Formation*, ed. by B. Dutta, and M. O. Jackson, pp. 97–140. Springer Verlag, New York, NY.

——— (2005a): "Allocation Rules for Network Games," *Games and Economic Behavior*, 51, 128–154.

——— (2005b): "A Survey of Models of Network Formation: Stability and Efficiency," in *Group Formation in Economics: Networks, Clubs, and Coalitions*, ed. by G. Demange, and M. Wooders, chap. 1, pp. 11–57. Cambridge University Press, Cambridge, United Kingdom.

JACKSON, M. O., AND A. WATTS (2001): "The Existence of Pairwise Stable Networks," *Seoul Journal of Economics*, 14(3), 299–321.

——— (2002a): "The Evolution of Social and Economic Networks," *Journal of Economic Theory*, 106, 265–295.

——— (2002b): "On the Formation of Interaction Networks in Social Coordination Games," *Games and Economic Behavior*, 41, 265–291.

JACKSON, M. O., AND A. WOLINSKY (1996): "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory*, 71, 44–74.

MONDERER, D., AND L. S. SHAPLEY (1996): "Potential Games," *Games and Economic Behavior*, 14, 124–143.

MYERSON, R. B. (1977): "Graphs and Cooperation in Games," *Mathematics of Operations Research*, 2, 225–229.

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

QIN, C.-Z. (1996): "Endogenous Formation of Cooperation Structures," *Journal of Economic Theory*, 69, 218–226.

ROSENTHAL, R. W. (1973): "A Class of Games Possessing Pure-Strategy Nash Equilibria," *International Journal of Game Theory*, 2, 65–67.

SHAPLEY, L. (1953): "A Value for n-Person Games," in *Contributiuons to the Theory of Games*, ed. by R. Luce, and A. Tucker, vol. II. Princeton University Press, Princeton, NJ.

——— (1971): "Cores of Convex Games," *International Journal of Game Theory*, 1, 11–26.

SLIKKER, M. (2000): "Decision Making and Cooperation Structures," Ph.D. thesis, Tilburg University, Tilburg, The Netherlands.

SLIKKER, M., R. P. GILLES, H. NORDE, AND S. TIJS (2005): "Directed Networks, Payoff Properties, and Hierarchy Formation," *Mathematical Social Sciences*, 49, 55–80, in press.

SRIVASTAVA, V., J. NEEL, J. HICKS, A. MACKENZIE, K. LAU, L. DASILVA, J. REED, AND R. GILLES (2005): "Application of Game Theory to Distributed MANET Algorithms," *IEEE Communications Surveys and Tutorials*, forthcoming.

UI, T. (2000): "A Shapley Value Representation of Potential Games," *Games and Economic Behavior*, 31, 121–135.

VOORNEVELD, M., AND H. NORDE (1997): "A Characterization of Ordinal Potential Games," *Games and Economic Behavior*, 19, 235–242.

Appendix: Proofs of the main results

Proof of Theorem 3.3

We begin by showing that Equal Bargaining Power is necessary for a potential to exist. We shall use it in the subsequent proof of Theorem 3.3.

Claim 1 *Let φ be some network payoff function. If the Consent Game Γ_φ corresponding to φ admits an exact potential, then φ satisfies Equal Bargaining Power.*

Proof. Let φ be a network payoff function such the the corresponding Consent game Γ_φ admits an exact potential. Specifically, let $Q : A \to \mathbb{R}^N$ be that exact potential.
Now, assume to the contrary that φ does not satisfy the Equal Bargaining Power property. Then there exists a network $\widehat{g}$ and a link $ij \in \widehat{g}$ such that

$$\varphi_i(\widehat{g}) - \varphi_i(\widehat{g} - ij) \neq \varphi_j(\widehat{g}) - \varphi_j(\widehat{g} - ij). \tag{16}$$

Consider the non-superfluous strategy profile $\widehat{l}$ that supports $\widehat{g}$. Further, let $\widetilde{l}$ be the non-superfluous strategy profile supporting $\widehat{g} - ij$. Then,

$$\begin{aligned}
\widetilde{l}_{ij} &= \widetilde{l}_{ji} = 0;\\
\widehat{l}_{ij} &= \widehat{l}_{ji} = 1;\\
\widetilde{l}_{ik} &= \widehat{l}_{ik} \text{ for } k \neq i;\\
\widetilde{l}_{jk} &= \widehat{l}_{jk} \text{ for } k \neq j;\\
\widetilde{l}_{km} &= \widehat{l}_{km} \text{ for } k, m \notin \{i, j\}.
\end{aligned}$$

Now,

$$\pi_i(\widehat{l}) - \pi_i(\widetilde{l}_i, \widehat{l}_{-i}) = \varphi_i(\widehat{g}) - \varphi_i(\widehat{g} - ij) = Q(\widehat{l}) - Q(\widetilde{l}_i, \widehat{l}_{-i}). \tag{17}$$

Also,

$$\pi_j(\widehat{l}) - \pi_j(\widetilde{l}_j, \widehat{l}_{-j}) = \varphi_j(\widehat{g}) - \varphi_j(\widehat{g} - ij) = Q(\widehat{l}) - Q(\widetilde{l}_j, \widehat{l}_{-j}). \tag{18}$$

From (16), (17) and (18), it now follows that

$$Q(\widetilde{l}_i, \widehat{l}_{-i}) \neq Q(\widetilde{l}_j, \widehat{l}_{-j}). \tag{19}$$

On the other hand, note that

$$(\widetilde{l}_i, \widetilde{l}_j, \widehat{l}_{-i,j}) = \widetilde{l}$$

and that

$$g(\widetilde{l}_i, \widehat{l}_{-i}) = g(\widetilde{l}_j, \widehat{l}_{-j}) = g(\widetilde{l}_i, \widetilde{l}_j, \widehat{l}_{-i,j}) = g(\widetilde{l}) = \widehat{g} - ij.$$

Therefore,

$$\begin{aligned}
\pi_i(\widetilde{l}_j, \widehat{l}_{-j}) &= \pi_i(\widetilde{l}_i, \widetilde{l}_j, \widehat{l}_{-i,j}) = \varphi_i(\widehat{g} - ij);\\
\pi_j(\widetilde{l}_i, \widehat{l}_{-i}) &= \pi_j(\widetilde{l}_i, \widetilde{l}_j, \widehat{l}_{-i,j}) = \varphi_j(\widehat{g} - ij),
\end{aligned}$$

and

$$\pi_i(\widetilde{l}_j, \widehat{l}_{-j}) - \pi_i(\widetilde{l}_i, \widetilde{l}_j, \widehat{l}_{-i,j}) = 0;$$
$$\pi_j(\widetilde{l}_i, \widehat{l}_{-i}) - \pi_j(\widetilde{l}_i, \widetilde{l}_j, \widehat{l}_{-i,j}) = 0,$$

which implies

$$Q(\widetilde{l}_j, \widehat{l}_{-j}) - Q(\widetilde{l}_i, \widetilde{l}_j, \widehat{l}_{-i,j}) = 0;$$
$$Q(\widetilde{l}_i, \widehat{l}_{-i}) - Q(\widetilde{l}_i, \widetilde{l}_j, \widehat{l}_{-i,j}) = 0.$$

Thus, we conclude that

$$Q(\widetilde{l}_j, \widehat{l}_{-j}) = Q(\widetilde{l}_i, \widehat{l}_{-i}) = Q(\widetilde{l}_i, \widetilde{l}_j, \widehat{l}_{-i,j}). \tag{20}$$

But (20) contradicts (19). Hence, we have shown the assertion. ■

With the proven claim we are able to construct a proof of Theorem 3.3.

(i) implies (ii)

Suppose there exists an exact network potential for φ. Let $\omega\colon \mathbb{G}^N \to \mathbb{R}$ denote this exact network potential. We now show that an exact potential exists for the corresponding the Consent Game Γ_φ.

Consider any arbitrary action-tuple $l = (l_i)_{i \in N}$. Let $g(l) = g$ (say).

Let $i \in N$. Now, $N_i(g)$, as defined, is the set of players in i's neighborhood. Let $M_i(g) = \{j \in N \backslash N_i(g) \mid l_{ji} = 1\}$. Now, by deviating from l, i can form links with any member of the set $M_i(g)$ and can delete links with any member of the set $N_i(g)$. Hence, consider any arbitrary strategy $\widehat{l}_i \neq l_i$. Let

$$\widehat{N}_i(g) = \{j \in N_i(g) \mid \widehat{l}_{ij} = 0\};$$
$$\widehat{M}_i(g) = \{j \in M_i(g) \mid \widehat{l}_{ij} = 1\}.$$

Then,

$$g(\widehat{l}_i, l_{-i}) = g + \Big(\{i\} \times \widehat{N}_i(g)\Big) - \Big(\{i\} \times \widehat{M}_i(g)\Big) \tag{21}$$

Let $\#\widehat{N}_i(g) = k$ and $\#\widehat{M}_i(g) = m$ where $k, m \in \mathbb{N}$. Label the players such that $\{i_1, i_2, \ldots, i_k\} = \widehat{N}_i(g)$ and $\{i_{k+1}, i_{k+2}, \ldots, i_{k+m}\} = \widehat{M}_i(g)$. Next construct a sequence of adjacent networks as follows:

$$\begin{aligned}
g_1 &= g - ii_1;\\
g_2 &= g_1 - ii_2;\\
&\vdots\\
g_k &= g_{k-1} - ii_k;\\
g_{k+1} &= g_k + ii_{k+1};\\
g_{k+2} &= g_{k+1} + ii_{k+2};\\
&\vdots\\
g_{k+m} &= g_{k+m-1} + ii_{k+m}.
\end{aligned}$$

From (21), it then follows that

$$g_{k+m} = g(\widehat{l}_i, l_{-i}). \tag{22}$$

Now, from (6), it follows that

$$\begin{aligned}
\omega(g) - \omega(g_1) &= \varphi_i(g) - \varphi_i(g_1); \\
\omega(g_1) - \omega(g_2) &= \varphi_i(g_1) - \varphi_i(g_2); \\
&\vdots \\
\omega(g_{k-1}) - \omega(g_k) &= \varphi_i(g_{k-1}) - \varphi_i(g_k).
\end{aligned}$$

Adding both sides of all the expressions (which results in common terms cancelling out), we get

$$\omega(g) - \omega(g_k) = \varphi_i(g) - \varphi_i(g_k). \tag{23}$$

Similarly, from (6), it follows that

$$\omega(g_{k+m}) - \omega(g_k) = \varphi_i(g_{k+m}) - \varphi_i(g_k). \tag{24}$$

Taking the difference of (23) and (24), we get

$$\omega(g_{k+m}) - \omega(g) = \varphi_i(g_{k+m}) - \varphi_i(g). \tag{25}$$

From (22) and (25), it follows

$$\omega(g(\widehat{l}_i, l_{-i})) - \omega(g(l)) = \varphi_i(g(\widehat{l}_i, l_{-i})) - \varphi_i(g(l)) = \pi_i(\widehat{l}_i, l_{-i}) - \pi_i(l). \tag{26}$$

Define a function $Q : A \to \mathbb{R}^N$ such that $Q(l) = \omega(g(l))$. From (26), it now follows that

$$Q(\widehat{l}_i, l_{-i}) - Q(l) = \pi_i(\widehat{l}_i, l_{-i}) - \pi_i(l).$$

This shows that, by definition, Q is indeed an exact potential of the corresponding Consent Game Γ_φ.

(ii) implies (i)
Suppose the Consent Game Γ_φ admits an exact potential and that this potential is given by $Q\colon A \to \mathbb{R}$. We will show that φ now admits an exact network potential. We know from Claim 1 that Equal Bargaining Power must be satisfied. Hence, for all $ij \in g$, $g \in \mathbb{G}^N$,

$$\varphi_i(g) - \varphi_i(g - ij) = \varphi_j(g) - \varphi_j(g - ij). \tag{27}$$

Define $\omega\colon \mathbb{G}^N \to \mathbb{R}$ such that

$$\omega(g) = Q(l_g) \tag{28}$$

where l_g is the non-superfluous strategy supporting g. We shall show ω is an exact network potential for φ.
Consider any arbitrary g and the corresponding l_g and any $ij \in g$. Let $\widehat{l}_{ij} = 0$ and $\widehat{l}_{ik} = l_{g,ik}$ for all $k \neq j$. From the definition of an exact potential we conclude that

$$Q(l_g) - Q(\widehat{l}_i, l_{g,-i}) = \pi_i(l_g) - \pi_i(\widehat{l}_i, l_{g,-i}) \tag{29}$$

Also,

$$\pi_i(l_g) - \pi_i(\widehat{l}_i, l_{g,-i}) = \varphi_i(g(l_g)) - \varphi_i(g(\widehat{l}_i, l_{g,-i})) = \varphi_i(g) - \varphi_i(g - ij). \tag{30}$$

From (28), (29) and (30), it then follows that

$$\omega(g) - Q(\widehat{l}_i, l_{g,-i}) = \varphi_i(g) - \varphi_i(g - ij). \tag{31}$$

Now, we claim that

$$Q(\widehat{l}_i, l_{g,-i}) = Q(l_{g-ij}) = \omega(g - ij). \tag{32}$$

We shall prove (32) as follows. First, by the definition of an exact potential,

$$\pi_j(\widehat{l}_i, l_{g,-i}) - \pi_j(\widehat{l}_i, \widehat{l}_j, l_{g,-i,j}) = Q(\widehat{l}_i, l_{g,-i}) - Q(\widehat{l}_i, \widehat{l}_j, l_{g,-i,j}) \tag{33}$$

where $\widehat{l}_j$ is such that $\widehat{l}_{ji} = 0$, $\widehat{l}_{jk} = l_{g,jk}$ for all $k \neq i$. Moreover,

$$\begin{aligned} \pi_j(\widehat{l}_i, l_{g,-i}) - \pi_j(\widehat{l}_i, \widehat{l}_j, l_{g,-i,j}) &= \varphi_j(g(\widehat{l}_i, l_{g,-i})) - \varphi_j(g(\widehat{l}_i, \widehat{l}_j, l_{g,-i,j})) \\ &= \varphi_j(g - ij) - \varphi_j(g - ij) = 0. \end{aligned} \tag{34}$$

Also,

$$\left(\widehat{l}_i, \widehat{l}_j, l_{g,-i,j}\right) = l_{g-ij}. \tag{35}$$

Hence, from (33), (34) and (35), we conclude that (32) indeed holds.
To conclude the proof of the assertion, we point out that (27), (31) and (32) implies that

$$\varphi_i(g) - \varphi_i(g - ij) = \varphi_j(g) - \varphi_j(g - ij) = \omega(g) - \omega(g - ij).$$

This proves that ω is indeed an exact network potential.

(i) implies (iii)
Consider any $g \in \mathbb{G}^N$ and let ρ_1 as well as ρ_2 be two arbitrary orders on g. Now suppose that there exists an exact network potential for φ, say ω. We will show that φ has to satisfy the Sums Property.
Assume without loss of generality that the network g consist of exactly k links. Hence, for order ρ_1 on g,

$$\begin{aligned} \omega(g_1^{\rho_1}) - \omega(g_0) &= \theta_\varphi(g_1^{\rho_1}, \tau_1^{\rho_1}); \\ \omega(g_2^{\rho_1}) - \omega(g_1^{\rho_1}) &= \theta_\varphi(g_2^{\rho_1}, \tau_2^{\rho_1}); \\ &\vdots \\ \omega(g) - \omega(g_{k-1}^{\rho_1}) &= \theta_\varphi(g_k^{\rho_1}, \tau_k^{\rho_1}). \end{aligned}$$

Adding both sides and cancelling out common terms, we get

$$\omega(g) - \omega(g_0) = S_\varphi(g, \rho_1). \tag{36}$$

Similarly, we conclude that

$$\omega(g) - \omega(g_0) = S_\varphi(g, \rho_2). \tag{37}$$

Hence, (36) and (37) imply that

$$S_\varphi(g, \rho_1) = S_\varphi(g, \rho_2).$$

Hence, we conclude that φ indeed satisfies the Sums Property.

(iii) implies (i)
Suppose that φ satisfies the Sums Property. Then for any arbitrary network $g \in \mathbb{G}^N$ consisting of k links and all orders $\rho_1, \rho_2, \ldots, \rho_T \in X^g$ with $T = k!$,

$$S_\varphi(g, \rho_1) = S_\varphi(g, \rho_2) = \ldots = S_\varphi(g, \rho_T) = S_\varphi(g) \text{ (say)}.$$

We proceed now to prove that the function $S_\varphi \colon \mathbb{G}^N \to \mathbb{R}$ is an exact network potential for φ.
Take any arbitrary network $g \in \mathbb{G}^N$ with k links and any arbitrary link $ij \in g$. Obviously, we can select the order $\rho_T \in X^g$ on network g such that $ij = \tau_k^{\rho_T}$. Consider the sequence of links $\tau_1^{\rho_T}, \tau_2^{\rho_T}, \ldots, \tau_{k-1}^{\rho_T}$. But this sequence of links constitutes an order on $g - ij$. Call this particular order $\rho'_{g-ij} \in X^{g-ij}$. Therefore,

$$\begin{aligned}
S_\varphi(g, \rho_T) &= S_\varphi(g) = \sum_{p=1}^{k} \theta_\varphi(g_p^{\rho_T}, \tau_p^{\rho_T}) \\
S_\varphi(g, \rho'_{g-ij}) &= S_\varphi(g - ij) = \sum_{p=1}^{k-1} \theta_\varphi(g_p^{\rho'_{g-ij}}, \tau_p^{\rho'_{g-ij}}) = \\
&= \sum_{p=1}^{k-1} \theta_\varphi(g_p^{\rho_T}, \tau_p^{\rho_T}).
\end{aligned}$$

Therefore,

$$\begin{aligned}
S_\varphi(g) - S_\varphi(g - ij) &= \theta_\varphi(g_k^{\rho_T}, \tau_k^{\rho_T}) \\
&= \theta_\varphi(g, ij) \\
&= \varphi_i(g) - \varphi_i(g - ij) \\
&= \varphi_j(g) - \varphi_j(g - ij).
\end{aligned}$$

This proves that $S_\varphi \colon \mathbb{G}^N \to \mathbb{R}$ is indeed an exact network potential for φ.

Proof of Theorem 3.7

In this section we proceed with the development of a proof of Theorem 3.7. The following two intermediary results are essential for our proof. The first one is the well-known four-cycle property for exact potential games.

Claim 2 (Monderer and Shapley 1996, Corollary 2.9)
The game (A, π) *has a potential if and only if for all* $i, j \in N$*, all* $\widehat{a}_{-i,j} \in \prod_{k \neq i,j} A_k$*, all* $a_i^1, a_i^2 \in A_i$ *and all* $a_j^1, a_j^2 \in A_j$*,*

$$\pi_i(\beta) - \pi_i(\alpha) + \pi_j(\gamma) - \pi_j(\beta) + \pi_i(\delta) - \pi_i(\gamma) + \pi_j(\alpha) - \pi_j(\delta) = 0$$

where

$$\begin{aligned}
\alpha &= (a_i^1, a_j^1, \widehat{a}_{-i,j}), \\
\beta &= (a_i^2, a_j^1, \widehat{a}_{-i,j}), \\
\gamma &= (a_i^2, a_j^2, \widehat{a}_{-i,j}), \\
\delta &= (a_i^1, a_j^2, \widehat{a}_{-i,j}).
\end{aligned}$$

Our proof of Theorem 3.7 mainly rests on the following new insight:

Claim 3 *Consider some network* $g \in \mathbb{G}^N$*. If* $i \in N_0(g)$ *and* $\varphi \in \widetilde{\Phi}$*, then* $\phi_i(U_{\varphi,g}) = \varphi_i(g)$*.*

Proof. Let $g \in \mathbb{G}^N$, $\varphi \in \widetilde{\Phi}$, and $i \in N_0(g)$ be as stated in the assertion.
We know from (3) that

$$\phi_i(U_{\varphi,g}) = \sum_{S \subset N:\, S \ni i} \frac{(|S|-1)!(n-|S|)!\,[U_{\varphi,g}(S) - U_{\varphi,g}(S \backslash \{i\})]}{n!}$$

where $U_{\varphi,g}(S) = \sum_{i \in S} \varphi_i(g|_S)$ and $g|_S = \{ij \in g \mid i \in S \text{ and } j \in S\}$.
Consider any coalition $S \subset N$ that includes i. Now, since $i \in N_0(g)$, we have that $g|_S = g|_{S \backslash \{i\}}$. Hence, the payoffs of all players are identical in both networks $g|_S$ and $g|_{S \backslash \{i\}}$ for any coalition S that includes i. Therefore,

$$\begin{aligned}
U_{\varphi,g}(S) &= \sum_{j \in S} \varphi_j(g|_S) \\
&= \varphi_i(g|_S) + \sum_{j \in S \backslash \{i\}} \varphi_j(g|_S) \\
&= \varphi_i(g|_{S \backslash \{i\}}) + \sum_{j \in S \backslash \{i\}} \varphi_j(g|_{S \backslash \{i\}}) \\
&= \varphi_i(g|_{S \backslash \{i\}}) + U_{\varphi,g}(S \backslash \{i\}).
\end{aligned}$$

Obviously, since $i \in N_0(g)$, it holds that $\varphi_i(g|_S) = \varphi_i(g|_{S \backslash \{i\}}) = \tau_i$. Hence,

$$U_{\varphi,g}(S) - U_{\varphi,g}(S \backslash \{i\}) = \tau_i.$$

This, in turn, implies that

$$\phi_i(U_{\varphi,g}) = \tau_i \cdot \sum_{S\subset N:\, S\ni i} \frac{(|S|-1)!(n-|S|)!}{n!} = \tau_i = \varphi_i(g).$$

This shows the claim. ■

Proof of Theorem 3.7(a).
For a given φ, construct a new payoff function $\widehat{\varphi}(g) = \phi(U_{\varphi,g})$ for all $g \in \mathbb{G}^N$. We construct the non-cooperative game $(A, \widehat{\pi})$ where $\widehat{\pi} = (\widehat{\pi}_i)_{i\in N} \colon A \to \mathbb{R}^N$ such that $\widehat{\pi}_i(l) = \widehat{\varphi}_i(g(l))$ for all $l \in A$. Hence, $(A, \widehat{\pi})$ is the Consent Game corresponding to $\widehat{\varphi}$. We shall first show that $(A, \widehat{\pi})$ has an exact potential, and hence by Theorem 3.3, $\widehat{\varphi}$ has an exact network potential. This implies that if φ is Shapley-consistent in the sense that $\varphi = \widehat{\varphi}$, then φ admits an exact network potential as well.
Now, we know from (4) that

$$\phi_i(U_{\varphi,g}) - \phi_j(U_{\varphi,g}) = \sum_{S\subset N\setminus\{ij\}} \frac{(|S|)!(n-|S|-2)!\,[U_{\varphi,g}(S\cup\{i\}) - U_{\varphi,g}(S\cup\{j\})]}{(n-1)!}$$

which implies that

$$\widehat{\varphi}_i(g) - \widehat{\varphi}_j(g) = \sum_{S\subset N\setminus\{ij\}} \frac{(|S|)!(n-|S|-2)!\,[U_{\varphi,g}(S\cup\{i\}) - U_{\varphi,g}(S\cup\{j\})]}{(n-1)!}.$$

Therefore for any strategy tuple $l \in A$ of the Consent Game $(A, \widehat{\pi})$,

$$\widehat{\varphi}_i(g(l)) - \widehat{\varphi}_j(g(l)) = \sum_{S\subset N\setminus\{ij\}} \frac{(|S|)!(n-|S|-2)!\left[U_{\varphi,g(l)}(S\cup\{i\}) - U_{\varphi,g(l)}(S\cup\{j\})\right]}{(n-1)!}$$

which implies

$$\begin{aligned}&\widehat{\pi}_i(l) - \widehat{\pi}_j(l) = \\ &\quad = \sum_{S\subset N\setminus\{ij\}} \frac{(|S|)!(n-|S|-2)!\left[U_{\varphi,g(l)}(S\cup\{i\}) - U_{\varphi,g(l)}(S\cup\{j\})\right]}{(n-1)!}.\end{aligned}$$

Consider any arbitrary $l_i^1, l_i^2 \in A_i$, $l_j^1, l_j^2 \in A_j$, and $\widehat{l}_{-i,j} \in \prod_{k\neq i,j} A_k$ and define

$$\begin{aligned}\alpha &= (l_i^1, l_j^1, \widehat{l}_{-i,j}),\\ \beta &= (l_i^2, l_j^1, \widehat{l}_{-i,j}),\\ \gamma &= (l_i^2, l_j^2, \widehat{l}_{-i,j}),\\ \delta &= (l_i^1, l_j^2, \widehat{l}_{-i,j}).\end{aligned}$$

Then,

$$
\begin{aligned}
& \widehat{\pi}_i(\beta) - \widehat{\pi}_i(\alpha) + \widehat{\pi}_j(\gamma) - \widehat{\pi}_j(\beta) + \widehat{\pi}_i(\delta) - \widehat{\pi}_i(\gamma) + \widehat{\pi}_j(\alpha) - \widehat{\pi}_j(\delta) = \\
= \; & [\widehat{\pi}_j(\alpha) - \widehat{\pi}_i(\alpha)] + [\widehat{\pi}_i(\beta) - \widehat{\pi}_j(\beta)] + [\widehat{\pi}_j(\gamma) - \widehat{\pi}_i(\gamma)] + [\widehat{\pi}_i(\delta) - \widehat{\pi}_j(\delta)] \\
= \; & \sum_{S \subset N \backslash \{ij\}} \frac{(|S|)!(n-|S|-2)!}{(n-1)!} \left[U_{\varphi,g(\alpha)}(S \cup \{j\}) - U_{\varphi,g(\alpha)}(S \cup \{i\}) \right] \\
& + \sum_{S \subset N \backslash \{ij\}} \frac{(|S|)!(n-|S|-2)!}{(n-1)!} \left[U_{\varphi,g(\beta)}(S \cup \{i\}) - U_{\varphi,g(\beta)}(S \cup \{j\}) \right] \\
& + \sum_{S \subset N \backslash \{ij\}} \frac{(|S|)!(n-|S|-2)!}{(n-1)!} \left[U_{\varphi,g(\gamma)}(S \cup \{j\}) - U_{\varphi,g(\gamma)}(S \cup \{i\}) \right] \\
& + \sum_{S \subset N \backslash \{ij\}} \frac{(|S|)!(n-|S|-2)!}{(n-1)!} \left[U_{\varphi,g(\delta)}(S \cup \{i\}) - U_{\varphi,g(\delta)}(S \cup \{j\}) \right] \\
= \; & \sum_{S \subset N \backslash \{ij\}} \frac{(|S|)!(n-|S|-2)!}{(n-1)!} \left[U_{\varphi,g(\alpha)}(S \cup \{j\}) - U_{\varphi,g(\beta)}(S \cup \{j\}) \right] \\
& + \sum_{S \subset N \backslash \{ij\}} \frac{(|S|)!(n-|S|-2)!}{(n-1)!} \left[U_{\varphi,g(\beta)}(S \cup \{i\}) - U_{\varphi,g(\gamma)}(S \cup \{i\}) \right] \\
& + \sum_{S \subset N \backslash \{ij\}} \frac{(|S|)!(n-|S|-2)!}{(n-1)!} \left[U_{\varphi,g(\gamma)}(S \cup \{j\}) - U_{\varphi,g(\delta)}(S \cup \{j\}) \right] \\
& + \sum_{S \subset N \backslash \{ij\}} \frac{(|S|)!(n-|S|-2)!}{(n-1)!} \left[U_{\varphi,g(\delta)}(S \cup \{i\}) - U_{\varphi,g(\alpha)}(S \cup \{i\}) \right].
\end{aligned}
$$

Now, we claim that each of the four expressions in square brackets are zero. Start with the first expression $U_{\varphi,g(\alpha)}(S \cup \{j\}) - U_{\varphi,g(\beta)}(S \cup \{j\})$ for any $S \subset N \backslash \{ij\}$. α and β only differ in that player i changes her strategy from l_i^1 in α to l_i^2 in β. All other players keep their strategies unchanged. But from the definition, both $U_{\varphi,g(\alpha)}(S \cup \{j\})$ and $U_{\varphi,g(\beta)}(S \cup \{j\})$ are independent of strategies chosen by i. Therefore, $U_{\varphi,g(\alpha)}(S \cup \{j\}) - U_{\varphi,g(\beta)}(S \cup \{j\}) = 0$. The same holds for the three other bracketed expressions as well. Hence, by Claim 2, $(A, \widehat{\pi})$ admits an exact potential.

Proof of Theorem 3.7(b).

Let the network payoff function φ admit an exact network potential. This implies by Theorem 3.3 that the corresponding Consent Game Γ_φ admits an exact potential. Let the potential function of the Consent Game corresponding to φ be denoted by P. Suppose to the contrary that the assertion does not hold and assume that $\varphi(g) \neq \phi(U_{\varphi,g})$ for some $g \in \mathbb{G}^N$.

Now, we know from the earlier proof that if the payoff function were $\widehat{\varphi}$ where $\widehat{\varphi}(g) = \phi(U_{\varphi,g})$ for all $g \in \mathbb{G}^N$, then also the corresponding Consent Game has an exact potential. Let the potential function of the Consent Game corresponding to $\widehat{\varphi}$ be denoted by $\widehat{P}$.

Claim: *There exists a constant* $C \in \mathbb{R}$ *such that* $P(l) = \widehat{P}(l) + C$ *for all* $l \in A$.

We prove the claim by induction. Let for any $i \in N$, $l_i \in A_i$, $\xi(l_i) = \#\{j \in N \mid l_{ij} = 1\}$. Define an order $\prec$ on A such that $l \prec l'$ if and only if there exists $i \in N$ such that $\xi(l_i) < \xi(l'_i)$ and for all $j < i$, $j \in N$, $\xi(l_j) = \xi(l'_j)$. Then, the smallest element of A with respect to $\prec$ is given by l^0 where $l^0_{ij} = 0$ for all $i, j \in N$, $i \neq j$. Set $C = \widehat{P}(l^0) - P(l^0)$. We shall show that if for all $l \prec l'$, it holds that $\widehat{P}(l) = P(l) + C$, then it is the case that $\widehat{P}(l') = P(l') + C$.
For every $i \in N$, define $l^i \in A$ as $l^i_{ij} = 0$ for all $j \neq i$ and $l^i_{jk} = l_{jk}$ for all $j \neq i$, $k \in N$. Then, $l^i \prec l$ and $i \in N_0(g(l^i))$.
Also, by definition of an exact potential, it must be the case that

$$\begin{aligned}
& \pi_i(l) - \pi_i(l^i) = P(l) - P(l^i) \\
\Longrightarrow \quad & \varphi_i(g(l)) - \varphi_i(g(l^i)) = P(l) - P(l^i) \\
\Longrightarrow \quad & \sum_{i \in N} \varphi_i(g(l)) - \sum_{i \in N} \varphi_i(g(l^i)) = n \cdot P(l) - \sum_{i \in N} P(l^i) \\
\Longrightarrow \quad & P(l) = \frac{\sum_{i \in N} \varphi_i(g(l)) - \sum_{i \in N} \varphi_i(g(l^i)) + \sum_{i \in N} P(l^i)}{n}
\end{aligned} \tag{38}$$

Similarly,

$$\widehat{P}(l) = \frac{\sum_{i \in N} \widehat{\varphi}_i(g(l)) - \sum_{i \in N} \widehat{\varphi}_i(g(l^i)) + \sum_{i \in N} \widehat{P}(l^i)}{n}. \tag{39}$$

Now, given that the Shapley value is a feasible allocation, we get that

$$\sum_{i \in N} \widehat{\varphi}_i(g(l)) = U_{\varphi, g(l)}(N) = \sum_{i \in N} \varphi_i(g(l)).$$

From Claim 3 and $i \in N_0(g(l^i))$, $\varphi_i(g(l^i)) = \widehat{\varphi}_i(g(l^i)) = \tau_i$. Taking the difference of (38) and (39), we get by the induction hypothesis that

$$P(l) - \widehat{P}(l) = \frac{1}{n} \sum_{i \in N} \left[P(l^i) - \widehat{P}(l^i) \right] = C.$$

Hence, we have shown the claim. $\square$

To proceed with the proof of Theorem 3.7(b), we remark that for any $l \in A$, $\pi_i(l) - \pi_i(l^i) = P(l) - P(l^i)$ and $\widehat{\pi}_i(l) - \widehat{\pi}_i(l^i) = \widehat{P}(l) - \widehat{P}(l^i)$. The claim shown above then implies that

$$\begin{aligned}
\left[\pi_i(l) - \pi_i(l^i)\right] - \left[\widehat{\pi}_i(l) - \widehat{\pi}_i(l^i)\right] &= \left[P(l) - P(l^i)\right] - \left[\widehat{P}(l) - \widehat{P}(l^i)\right] \\
&= \left[P(l) - \widehat{P}(l)\right] - \left[P(l^i) - \widehat{P}(l^i)\right] \\
&= C - C = 0.
\end{aligned}$$

Hence, for any $l \in A$,

$$\pi_i(l) - \widehat{\pi}_i(l) = \pi_i(l^i) - \widehat{\pi}_i(l^i). \tag{40}$$

But we know from $i \in N_0(g(l^i))$ that

$$\pi_i(l^i) - \widehat{\pi}_i(l^i) = \varphi_i(g(l^i) - \widehat{\varphi}_i(g(l^i) = \tau_i - \tau_i = 0$$

which implies using (40) that

$$\pi_i(l) = \widehat{\pi}_i(l)$$

for any $l \in A$. Hence, $\varphi = \widehat{\varphi}$ which is a contradiction. This shows the assertion.

Proof of Theorem 4.3

We require the following intermediary result. It states that PSC is necessary for an ordinal potential to exist.

Claim 4 *If the Consent Game Γ_φ admits an ordinal potential, then the corresponding network payoff function φ satisfies PSC.*

Proof. Let Γ_φ admit an ordinal potential. Denote this potential function as $Q\colon A \to \mathbb{R}^N$. By contradiction, now assume that φ violates PSC. Then, there exists a network $\widehat{g}$ and a link $ij \in \widehat{g}$ such that either of the following three conditions hold:

$$\varphi_i(\widehat{g}) - \varphi_i(\widehat{g} - ij) > 0 \quad \text{and} \quad \varphi_j(\widehat{g}) - \varphi_j(\widehat{g} - ij) \leqslant 0; \tag{41}$$
$$\varphi_i(\widehat{g}) - \varphi_i(\widehat{g} - ij) < 0 \quad \text{and} \quad \varphi_j(\widehat{g}) - \varphi_j(\widehat{g} - ij) \geqslant 0; \tag{42}$$
$$\varphi_i(g) - \varphi_i(g - ij) = 0 \quad \text{and} \quad \varphi_j(g) - \varphi_j(g - ij) \neq 0. \tag{43}$$

We shall show that each of these requirements leads to a contradiction. Consider the non-superfluous strategy profile $\widehat{l}$ that supports $\widehat{g}$. Further, let $\widetilde{l}$ be the non-superfluous strategy profile supporting $\widehat{g} - ij$. Then,

$$\begin{aligned}
&\widetilde{l}_{ij} = \widetilde{l}_{ji} = 0;\\
&\widehat{l}_{ij} = \widehat{l}_{ji} = 1;\\
&\widetilde{l}_{ik} = \widehat{l}_{ik} \text{ for } k \neq i;\\
&\widetilde{l}_{jk} = \widehat{l}_{jk} \text{ for } k \neq j;\\
&\widetilde{l}_{kl} = \widehat{l}_{kl} \text{ for } k \neq i, j.
\end{aligned}$$

Note that

$$\left(\widetilde{l}_i, \widetilde{l}_j, \widehat{l}_{-i,j}\right) = \widetilde{l}.$$

Also,

$$g(\widetilde{l}_i, \widehat{l}_{-i}) = g(\widetilde{l}_j, \widehat{l}_{-j}) = g(\widetilde{l}_i, \widetilde{l}_j, \widehat{l}_{-i,j}) = g(\widetilde{l}) = \widehat{g} - ij.$$

Therefore,

$$\begin{aligned}
&\pi_i(\widetilde{l}_j, \widehat{l}_{-j}) = \pi_i(\widetilde{l}_i, \widetilde{l}_j, \widehat{l}_{-i,j}) = \varphi_i(\widehat{g} - ij);\\
&\pi_j(\widetilde{l}_i, \widehat{l}_{-i}) = \pi_j(\widetilde{l}_i, \widetilde{l}_j, \widehat{l}_{-i,j}) = \varphi_j(\widehat{g} - ij),
\end{aligned}$$

and

$$\pi_i(\widetilde{l}_j, \widehat{l}_{-j}) - \pi_i(\widetilde{l}_i, \widetilde{l}_j, \widehat{l}_{-i,j}) = 0;$$
$$\pi_j(\widetilde{l}_i, \widehat{l}_{-i}) - \pi_j(\widetilde{l}_i, \widetilde{l}_j, \widehat{l}_{-i,j}) = 0;$$

which by definition of Q as an ordinal potential implies that

$$Q(\widetilde{l}_j, \widehat{l}_{-j}) - Q(\widetilde{l}_i, \widetilde{l}_j, \widehat{l}_{-i,j}) = 0;$$
$$Q(\widetilde{l}_i, \widehat{l}_{-i}) - Q(\widetilde{l}_i, \widetilde{l}_j, \widehat{l}_{-i,j}) = 0;$$

This in turn implies that

$$Q(\widetilde{l}_j, \widehat{l}_{-j}) = Q(\widetilde{l}_i, \widehat{l}_{-i}) = Q(\widetilde{l}_i, \widetilde{l}_j, \widehat{l}_{-i,j}). \tag{44}$$

Now, suppose (41) holds. Then,

$$\pi_i(\widehat{l}) - \pi_i(\widetilde{l}_i, \widehat{l}_{-i}) = \varphi_i(\widehat{g}) - \varphi_i(\widehat{g} - ij) > 0$$

which by definition of Q as an ordinal potential implies that

$$Q(\widehat{l}) - Q(\widetilde{l}_i, \widehat{l}_{-i}) > 0. \tag{45}$$

Also,

$$\pi_j(\widehat{l}) - \pi_j(\widetilde{l}_j, \widehat{l}_{-j}) = \varphi_j(\widehat{g}) - \varphi_j(\widehat{g} - ij) \leqslant 0$$

which by definition of Q as an ordinal potential implies that

$$Q(\widehat{l}) - Q(\widetilde{l}_j, \widehat{l}_{-j}) \leqslant 0. \tag{46}$$

From (45) and (46), it now follows that

$$Q(\widetilde{l}_i, \widehat{l}_{-i}) \neq Q(\widetilde{l}_j, \widehat{l}_{-j}). \tag{47}$$

But (47) contradicts (44). Hence, we have a contradiction.

It should also be clear that if (42) holds, one can construct a similar proof that leads us to the desired contradiction.

Finally, if (43) holds,

$$\pi_i(\widehat{l}) - \pi_i(\widetilde{l}_i, \widehat{l}_{-i}) = \varphi_i(\widehat{g}) - \varphi_i(\widehat{g} - ij) = 0$$

which implies by definition of Q as an ordinal potential,

$$Q(\widehat{l}) - Q(\widetilde{l}_i, \widehat{l}_{-i}) = 0. \tag{48}$$

Also, either

$$\pi_j(\widehat{l}) - \pi_j(\widetilde{l}_j, \widehat{l}_{-j}) = \varphi_j(\widehat{g}) - \varphi_j(\widehat{g} - ij) > 0$$

in which case, from the definition of Q as an ordinal potential,

$$Q(\widehat{l}) - Q(\widetilde{l}_j, \widehat{l}_{-j}) > 0, \tag{49}$$

or

$$\pi_j(\widehat{l}) - \pi_j(\widetilde{l}_j, \widehat{l}_{-j}) = \varphi_j(\widehat{g}) - \varphi_j(\widehat{g} - ij) < 0$$

in which case, from the definition of Q as an ordinal potential,

$$Q(\widehat{l}) - Q(\widetilde{l}_j, \widehat{l}_{-j}) < 0. \tag{50}$$

In any case, (49) and (50) imply that

$$Q(\widehat{l}) - Q(\widetilde{l}_j, \widehat{l}_{-j}) \neq 0. \tag{51}$$

Now from (48) and (51),

$$Q(\widetilde{l}_i, \widehat{l}_{-i}) \neq Q(\widetilde{l}_j, \widehat{l}_{-j}).$$

which contradicts (44). Hence, we have arrived at a contradiction. This completes the proof of Claim 4. ■

Finally, we are able to construct the proof of Theorem 4.3.

Proof of Theorem 4.3.
Suppose an ordinal potential exists for the Consent Game Γ_φ, say $Q\colon A \to \mathbb{R}$. We will show that there exists an ordinal network potential for the corresponding network payoff function φ. We know from Claim 4 that PSC must be satisfied. Hence, for all $ij \in g$, $g \in \mathbb{G}^N$,

$$\text{Sign}\,(\varphi_i(g), \varphi_i(g - ij)) = \text{Sign}\,(\varphi_j(g), \varphi_j(g - ij)). \tag{52}$$

Define $\omega : \mathbb{G}^N \to \mathbb{R}$ by

$$\omega(g) = Q(l_g) \tag{53}$$

where l_g is the non-superfluous strategy supporting g. We shall show ω is an ordinal network potential corresponding to φ.
Consider any arbitrary g and the corresponding l_g and any $ij \in g$. From the definition of ordinal potentials,

$$\text{Sign}\,(Q(l_g), Q(\widehat{l}_i, l_{g,-i})) = \text{Sign}\,(\pi_i(l_g), \pi_i(\widehat{l}_i, l_{g,-i})) \tag{54}$$

where $\widehat{l}_{ij} = 0$ and $\widehat{l}_{ik} = l_{g,ik}$ for all $k \neq j$. But

$$\pi_i(l_g) - \pi_i(\widehat{l}_i, l_{g,-i}) = \varphi_i(g(l_g)) - \varphi_i(g(\widehat{l}_i, l_{g,-i})) = \varphi_i(g) - \varphi_i(g - ij). \tag{55}$$

From (53), (54) and (55), it follows that

$$\text{Sign}\,(\omega(g), Q(\widehat{l}_i, l_{g,-i})) = \text{Sign}\,(\varphi_i(g), \varphi_i(g - ij)). \tag{56}$$

Now, we claim that

$$Q(\widehat{l}_i, l_{g,-i}) = Q(l_{g-ij}) = \omega(g - ij). \tag{57}$$

We show this assertion as follows. First, by definition of ordinal potentials,

$$\text{Sign}\,(\pi_j(\widehat{l}_i, l_{g,-i}), \pi_j(\widehat{l}_i, \widehat{l}_j, l_{g,-i,j})) = \text{Sign}\,(Q(\widehat{l}_i, l_{g,-i}), Q(\widehat{l}_i, \widehat{l}_j, l_{g,-i,j})) \tag{58}$$

where $\widehat{l}_j$ is such that $\widehat{l}_{ji} = 0$, $\widehat{l}_{jk} = l_{g,jk}$ for all $k \neq i$. However,

$$\begin{aligned}\pi_j(\widehat{l}_i, l_{g,-i}) - \pi_j(\widehat{l}_i, \widehat{l}_j, l_{g,-i,j}) &= \varphi_j(g(\widehat{l}_i, l_{g,-i})) - \varphi_j(g(\widehat{l}_i, \widehat{l}_j, l_{g,-i,j})) \\ &= \varphi_j(g - ij) - \varphi_j(g - ij) = 0. \end{aligned} \tag{59}$$

Also,

$$\left(\widehat{l}_i, \widehat{l}_j, l_{g,-i,j}\right) = l_{g-ij}. \tag{60}$$

Hence, by (58), (59), (60) and definition of ω, it follows that

$$\begin{aligned} & \text{Sign}\,(\pi_j(\widehat{l}_i, l_{g,-i}), \pi_j(\widehat{l}_i, \widehat{l}_j, l_{g,-i,j})) = 0 \\ \Longrightarrow\; & \text{Sign}\,(Q(\widehat{l}_i, l_{g,-i}), Q(\widehat{l}_i, \widehat{l}_j, l_{g,-i,j})) = 0 \\ \Longrightarrow\; & Q(\widehat{l}_i, l_{g,-i}) = Q(\widehat{l}_i, \widehat{l}_j, l_{g,-i,j}) \\ \Longrightarrow\; & Q(\widehat{l}_i, l_{g,-i}) = Q(l_{g-ij}) = \omega(g - ij) \end{aligned} \tag{61}$$

which proves the claim.
Finally, from (52), (56) and (57), it follows that

$$\text{Sign}\,(\varphi_i(g), \varphi_i(g - ij)) = \text{Sign}\,(\varphi_j(g), \varphi_j(g - ij)) = \text{Sign}\,(\omega(g), \omega(g - ij)).$$

This proves ω is indeed an ordinal network potential for φ. ■

Proof of Theorem 5.3

Claim 5 (Jackson and Watts 2001, Theorem 1A)
If there exists a function $w\colon \mathbb{G}^N \to \mathbb{R}$ such that $[g'$ defeats $g] \iff [w(g') > w(g)$ and g and g' are adjacent$]$, then there are no improvement cycles.

With this we are able to show assertion 5.3(b).

Claim 6 *If $\varphi \in \Phi_3$, then there are no improvement cycles.*

Proof. We have to find a function w as defined in Claim 5. But such a function is given by the ordinal network potential ω. Hence, the result follows. ■

From this assertion we immediately arrive on assertion 5.3(a) as a corollary to Claim 6.

Claim 7 *If $\varphi \in \Phi_3$, then there exists at least one pairwise stable network for the corresponding network payoff function.*

Proof. Follows from Lemma 5.2 and Claim 6. ∎

With regard to assertion 5.3(c) we remark that PSC is sufficient to guarantee that all strongly pairwise stable networks are strictly pairwise stable as well. Hence, we immediately establish the desired assertion.
This completes the proof of all three assertions stated in Theorem 5.3.

Proof of Theorem 5.5

In order to show Theorem 5.5, we use another result due to Jackson and Watts (2001).

Claim 8 (Jackson and Watts 2001, Theorem 1B)
If the network payoff function satisfies no indifference and there are no improvement cycles, then there exists a function $w\colon \mathbb{G}^N \to \mathbb{R}$ such that [g' defeats g] $\Leftrightarrow$ [$w(g) > w(g')$ and g and g' are adjacent].

To proceed with the proof of Theorem 5.5 we remark that by Claim 8, it follows that since the network payoff function satisfies no indifference and there are no improvement cycles, there exists a function $w : \mathbb{G}^N \to \mathbb{R}$ such that [g' defeats g] $\Leftrightarrow$ [$w(g') > w(g)$ and g and g' are adjacent].

Suppose g' defeats g. Then, it is either the case that

(i) $g' = g - ij$ with $\varphi_i(g') > \varphi_i(g)$, or

(ii) $g' = g + ij$ with $\varphi_i(g') > \varphi_i(g)$and $\varphi_j(g') \geqslant \varphi_j(g)$.

In both (i) and (ii), by PSC, $\varphi_j(g') > \varphi_j(g)$.
Hence, $[\,w(g) - w(g - ij) > 0\,] \Leftrightarrow [\,g$ defeats $g - ij\,] \Leftrightarrow [\,\varphi_i(g) - \varphi_i(g - ij) > 0$ and $\varphi_j(g) - \varphi_j(g - ij) > 0\,]$.
Also, $[\,w(g) - w(g - ij) < 0\,] \Leftrightarrow [\,g - ij$ defeats $g\,] \Leftrightarrow [\,\varphi_i(g) - \varphi_i(g - ij) < 0$ and $\varphi_j(g) - \varphi_j(g - ij) < 0\,]$.

Next, let $w(g) - w(g - ij) = 0$. Then neither $g - ij$ defeats g nor g defeats $g - ij$ which is impossible by no indifference. Also, if $\varphi_i(g) - \varphi_i(g - ij) = 0$ and $\varphi_j(g) - \varphi_j(g - ij) = 0$, again neither $g - ij$ defeats g nor g defeats $g - ij$ which is impossible by no indifference.
So it is trivially the case that $[\,w(g) - w(g - ij) = 0\,] \Leftrightarrow [\,\varphi_i(g) - \varphi_i(g - ij) = 0$ and $\varphi_j(g) - \varphi_j(g - ij) = 0\,]$.

Hence, by definition, w is an ordinal network potential, showing the assertion.

Proof of Theorem 5.6

For network payoff functions belonging to Φ_2, we prove that a strongly pairwise stable (and hence strictly pairwise stable) network exists by identifying those networks that are supported through ordinal potential maximizing strategies in Γ_φ. We show that these particular networks form a strict subset of the set of strongly pairwise stable networks $\mathcal{P}_s(\varphi) = \mathcal{P}^*(\varphi)$.

Claim 9 *Let $\varphi \in \Phi_2$ be such that the Consent Game $\Gamma_\varphi = (A, \pi)$ admits an ordinal potential function $Q \colon A \to \mathbb{R}$. if $l \in A$ is such that $Q(l) \geqq Q(l')$ for all $l' \in A$, then $g(l) \in \mathcal{P}_s(\varphi)$.*

Proof. Suppose that $g = g(l)$ is such that $l \in A$ maximizes the ordinal potential function Q on A. Now, assume by contradiction that g is not strongly pairwise stable. Then it is either not LAP or not SLDP. We shall show that both lead to contradictions.

First assume that g is not SLDP. This implies there exists a player $i \in N$ and a set of its neighbors $M \subset N_i(g)$ such that $\varphi_i(g \backslash h_M) > \varphi_i(g)$ where $h_M = \{ij \in g | j \in M\} \subset g$. Consider a strategy $\widehat{l}_i$ such that $\widehat{l}_{ij} = 0$ for all $j \in M$ and $\widehat{l}_{ij} = l_{ij}$ for all $j \in N \backslash M$. Then, $g(\widehat{l}_i, l_{-i}) = g \backslash h_M$. This implies $\pi_i(\widehat{l}_i, l_{-i}) > \pi_i(l)$, which in turn implies that $Q(\widehat{l}_i, l_{-i}) > Q(l)$. This contradicts the fact that l is a potential maximizer.

Next lets assume that g is not LAP. Then, there exists a pair $i, j \in N$ such that $\varphi_i(g + ij) > \varphi_i(g)$ implies $\varphi_j(g + ij) \geqslant \varphi_j(g)$. By Lemma 4, $\varphi_j(g + ij) > \varphi_j(g)$. Now, given that $ij \notin g$, there are three possibilities which we delineate as three cases. For each of these cases we show a contradiction.

CASE 1: $l_{ij} = 1$, $l_{ji} = 0$
In this case, consider a strategy $\widehat{l}_j$ such that $\widehat{l}_{ji} = 1$, $\widehat{l}_{jk} = l_{jk}$ for all $k \neq i$. Then, $\varphi_j(g + ij) > \varphi_j(g) \Rightarrow \pi_j(\widehat{l}_j, l_{-j}) > \pi_j(l) \Longrightarrow Q(\widehat{l}_j, l_{-j}) > Q(l)$ by definition of ordinal potentials which is a contradiction.

CASE 2: $l_{ij} = 0$, $l_{ji} = 1$
In this case, consider a strategy $\widehat{l}_i$ such that $\widehat{l}_{ij} = 1$, $\widehat{l}_{ik} = l_{ik}$ for all $k \neq j$. Then, $\varphi_i(g + ij) > \varphi_i(g) \Rightarrow \pi_i(\widehat{l}_i, l_{-i}) > \pi_i(l) \Longrightarrow Q(\widehat{l}_i, l_{-i}) > Q(l)$ by definition of ordinal potentials which is a contradiction.

CASE 3: $l_{ij} = 0 = l_{ji}$
Consider a strategy $\widehat{l}_i$ such that $\widehat{l}_{ij} = 1$, $\widehat{l}_{ik} = l_{ik}$ for all $k \neq j$. Then, $\pi_i(\widehat{l}_i, l_{-i}) = \pi_i(l) = \varphi_i(g) \Rightarrow Q(\widehat{l}_i, l_{-i}) = Q(l)$. Consider another strategy $\widehat{l}_j$ such that $\widehat{l}_{ji} = 1$, $\widehat{l}_{jk} = l_{jk}$ for all $k \neq i$. Then, $\pi_j(\widehat{l}_i, \widehat{l}_j, l_{-i,j}) = \varphi_j(g + ij) > \varphi_j(g) = \pi_j(\widehat{l}_i, l_{-i}) \Rightarrow Q(\widehat{l}_i, \widehat{l}_j, l_{-i,j}) > Q(\widehat{l}_i, l_{-i}) = Q(l)$ which contradicts that l is a potential maximizer.

That completes the proof of the claim. ∎