Stability of Networks with Continuous Stream of Payoffs: A Theory and Experiment*

Mariya Teteryatnikova†       James Tremewan‡

April 20, 2015

Abstract

We propose a new notion of farsighted pairwise stability for dynamic network formation which includes two notable features: cautiousness and consideration of intermediate payoffs. This differs from existing concepts which require a certain optimism on the part of the players in any environment without full communication and commitment, and which typically consider either only immediate or final payoffs. In a laboratory experiment we find some support for existing concepts, however our’s is the only one that identifies precisely the set of empirically stable networks.

KEY WORDS: networks, farsighted and cautious players, stability, improving path, experiment

JEL CLASSIFICATION: A14, C71, C92, D85

---

*We would like to thank Karl Schlag, Jean-Robert Tyran, Wieland Müller, Antonio Cabrales, Martin Cripps, Marco Mantovani, Nikita Roketskiy, Konrad Mierendorff, Philippe Jehiel, Antonio Guarino, Aureo de Paula, Aniol Llorente-Saguer, Asen Ivanov for useful suggestions and feedback. We also thank participants of seminars at the University of Vienna, University College London, School of Economics and Finance at Queen Mary University of London, Higher School of Economics and 2014 ESA European meeting for helpful comments. We acknowledge financial support from the Vienna Science and Technology Fund (WWTF) under project fund MA 09-017.

†University of Vienna. Email: mariya.teteryatnikova@univie.ac.at.

‡University of Vienna. Email: james.tremewan@univie.ac.at.
1 Introduction

In this paper we introduce a farsighted pairwise stability concept for dynamic network formation which includes two features not present in the most prominent existing definitions. Firstly, in contrast to many existing approaches to farsighted behaviour which require optimistic beliefs on the part of the players in any environment without full communication and commitment, we assume that players act to avoid any possibility of ending up worse off than in the status quo. Secondly, we assume that players may be interested not only in immediate payoffs (as in myopic concepts) or final payoffs (as in existing farsighted concepts) but also in payoffs accrued from intermediate steps. In a laboratory experiment we implement two network games designed to allow us to test and compare the predictive power of our new and existing concepts.

A number of farsighted pairwise stability concepts have been proposed in the literature. Most of these (pairwise farsightedly stable set, von Neumann-Morgenstern pairwise farsightedly stable set, largest pairwise consistent set (Herings et al., 2009; von Neumann and Morgenstern, 1944; Chwe, 1994), and level-K farsightedly stable set (Herings et al., 2014)) assume that players will often not remain in a network if there is a possibility of ending up better off as an eventual result of adding or deleting a link. However, there are instances where actually ending up in the desired network requires either good fortune, or full-communication and commitment. For example, after a first player deletes a link, a second player may have an equal incentive to delete either one of two further links to reach a stable network. Deleting one of these links makes the first player better off, but deleting the other makes them worse off. Under the aforementioned concepts, the current network is not stable because the first player may end up better off by deleting a link. However, if no credible commitment can be made by the second player to delete the “correct” link, it is reasonable to think that the first player may not be willing to take the risk, making the current network stable.

In our new concept, we assume that at least one of full communication or commitment is not possible.

---

1In this paper we focus on cooperative pairwise stability and do not consider the alternative approaches of either explicitly modeling a game and using non-cooperative equilibrium concepts or considering deviating coalitions of more than two players. Examples of the former include Myerson (1991), Jackson and Watts (2002b), Bala and Goyal (2000), Hojman and Szeidl (2008), Bloch (1996), Currarini and Morelli (2000), Goeree et al. (2009), Galeotti and Goyal (2010). Examples of the latter, with considerations of farsightedness in network formation, include Aumann and Mayerson (1988), Chwe (1994), Xue (1998), Dutta and Mutuswami (1997), Page Jr et al. (2005), Page and Wooders (2009), Herings et al. (2004), Mauleon and Vannetelbosch (2004).

2In coalition formation literature, the idea of cautiousness is not new. For example, Xue (1998) and Mauleon and Vannetelbosch (2004) consider pessimistic (or optimistic) attitudes of players to risk. One drawback of these approaches is that the proposed solution sets may be empty.
Furthermore, in the spirit of max-min strategies, we assume that players will not add or delete a link if there is any possibility that it will make them worse off in the long run.\(^3\)

Another common feature of these existing concepts is the assumption that players care only about the payoffs derived from the final (stable) network. At the other extreme are pairwise stability (Jackson and Wolinsky, 1996), pairwise myopic stability (Watts, 1997; Jackson and Watts, 2002a; Herings et al., 2009) and their refinements (Jackson and van den Nouweland, 2005; Belleflamme and Bloch, 2004; Goyal and Joshi, 2006), where the players consider only the payoffs that may be \textit{immediately} obtained from deleting or adding a link. However, in many applications agents may be also concerned with intermediate payoffs, and derive benefits or losses from their ongoing relationships. This is particularly common in the environments that are predisposed to exogenous shocks or frequent changes, for example, in interactions between banks in financial networks, between buyers and sellers in trading networks, between people in romance and acquaintance networks, between firms in business and collaboration networks. Our concept takes this into account, by allowing for arbitrary weighting of intermediate payoffs, including as special cases myopia and placing weight only on the final network.

One other paper that adopts a similar approach to modeling players’ payoffs is Dutta et al. (2005). It considers a dynamic network formation where individuals are farsighted and form or delete links so as to maximize the entire discounted stream of payoffs. Contrary to ours, this paper is closer in spirit to non-cooperative game theoretic models and proposes a specific protocol of network formation process.\(^4\) In fact, the properties of the \textit{process} of network formation, and not just the outcomes, are the key points of interest in the paper, as the main questions addressed concern the existence of a Markovian equilibrium process of network formation and its convergence to an efficient network. These features of the model – the structured approach to network formation and the focus on properties of the process of network formation, rather than just the outcomes, – are the principal distinctions from our paper.

In our paper, we propose a pairwise cooperative game theoretic approach to network formation and focus on the question of network stability in the environment where players are farsighted, cautious and

\(^3\)Such “extreme pessimism” is typical for a behaviour based on max-min type of preferences. We therefore expect our concept to perform best in the environments without very large differences in payoffs. We choose this approach since extreme pessimism is the simplest way to capture cautiousness in players’ behaviour, without having to deal with beliefs and weighting of many (or infinitely many) different alternatives. We note that this approach results in a larger set of stable networks than the alternative approaches, considering the weighted average. That is, networks which are not stable according to our definition, cannot be stable according to these alternative definitions.

\(^4\)The precisely defined rules of network formation process include a random choice of one pair of active players in every period, infinite horizon of the game, players following Markov strategies and taking into account the probability distribution over the feasible set of future networks, given the current state of the network, players’ strategies and the randomness of the active pair selection.
may care not only about their immediate or long-run payoffs but also about payoffs at intermediate steps of network formation. In the tradition of pairwise network formation literature, formation of a link requires a consent of both involved players, while severance of a link is a unilateral decision of any player involved in the link.\(^5\) By adding and deleting links with each other, players can consecutively transform the network, and a sequence of networks that emerge at each step of this transformation produces a so-called *path* between the initial and final network. We define two types of such paths, which then allows us to introduce our new stability concept. We call a path between two networks improving if all players involved in link changes on this path increase their payoffs relative to staying in the status quo network, and we call an improving path *surely* improving (relative to a set of networks \(G\)) if players’ payoffs increase not only on this path but also on *any credible* improving path that can be followed after the link change.\(^6\)

We note that our definition of the improving path includes as special cases the myopic and farsighted improving paths defined in Herings et al. (2009). These special cases arise when players derive payoffs only from the first or the last network of the path, respectively. More generally, when players also care about their intermediate payoffs, an improving path in our definition increases players’ payoffs associated with the path rather than the network, that is, players evaluate the improvements which the path offers relative to staying in the status quo network for the same number of steps. Moreover, the definition of the surely improving path incorporates the consideration that players are cautious and hence, will only follow the paths that increase their payoffs “with certainty”.

Given the above definitions, we introduce the concept of the *cautious path stable set* of networks. A set of networks \(G\) is cautious path stable if (1) for any pair of networks in the set, there is no surely improving path (relative to \(G\)) between them, (2) from any network outside the set, there exists a surely improving path (relative to \(G\)) leading to some network in the set, and (3) no proper subset of \(G\) satisfies conditions (1) and (2). In particular, a set consisting of a single network \(g\) is cautious path stable if there exists a surely improving path from any other network to \(g\). Conditions (1) and (2) of the definition impose internal and external stability. Alternatively, they can be viewed as the requirement that networks within a stable set are robust to perturbations, which might result from mistakes on the

---

\(^5\)Note that this aspect of only *partial* cooperation between the two players involved in a link establishes an important distinction between cooperative pairwise stability and coalitional stability. That is, while in the pairwise approach, only special 2-player “coalitions” can form, the cooperation in such coalitions is only partial, and every player has a natural “unilateral” domain of action. See discussion in Dutta et al. (2005).

\(^6\)The credibility of a path is determined by set \(G\), in the sense that only those improving paths that lead to \(G\) are deemed credible. Set \(G\) itself represents a stable set that we define below.
part of the players or some exogenous forces affecting the network.

Our definition of the cautious path stable set is similar to the definition of the von Neumann-Morgenstern pairwise farsightedly stable set (Herings et al., 2009; von Neumann and Morgenstern, 1944), but in contrast to the latter, it incorporates the consideration of intermediate payoffs and cautiousness in players behaviour when adding and deleting links. Moreover, in the special case when players only care about their end-of-path payoffs, this definition turns out to be close to the definition of the pairwise farsightedly stable set of Herings et al. (2009). In fact, we show that the same three conditions – deterrence of external deviations, external stability and minimality – are satisfied by our stable set, but in contrast to the concept of Herings et al. (2009), the external stability in our definition requires the existence of not just an improving but surely improving path from any network outside the set leading to some network in the set. Therefore, once again, players in our setting are more cautious.

We show that a cautious path stable set of networks always exists. We also provide some easy to verify conditions for a set to be cautious path stable and the unique cautious path stable set. Furthermore, in the setting where players care only about their end-of-path payoffs, we study the relationship between cautious path stable sets\(^7\) and sets identified as stable by some other farsighted stability concepts. We find that any cautious path stable set contains at least one pairwise farsightedly stable set (PWFS) as a subset. The converse – the inclusion of any PWFS set in some cautious path stable set, – is not necessarily true. However, a simple corollary of this statement is that if a PWFS set is unique, in which case it is also the unique von Neumann-Morgenstern pairwise farsightedly stable set (vN-MFS), then it is a subset of any cautious path stable set. Moreover, we find that if a cautious path stable set of networks satisfies an additional constraint, then it is a PWFS and a vN-MFS set. An even stricter constraint implies that a cautious path stable set is the unique PWFS and vN-MFS set. In particular, a cautious path stable set consisting of a single network is always PWFS and vN-MFS, and it is the unique stable set whenever no improving paths start at this network.

To provide a test of our new concept and compare it with other concepts we run a laboratory experiment. Such a test is useful as the concepts we are testing are general, and should apply equally in the laboratory as in the outside world. Thus, if a concept is not useful in explaining behaviour in the simplified laboratory environment, it is also unlikely to have predictive power in “real life” applications. In addition, we have the usual advantages of experimental over field studies in knowing the precise

---

\(^7\)In this setting, with the end-of-path payoff specification, we will later refer to our concept as a cautious final-network stable set.
payoffs associated with each network, and statistically independent repetitions of the same interactions.

In our laboratory experiment, we implement a “pure” network formation game where payoffs are derived directly from the network structure rather than from further interactions with linked players. We focus on these kinds of games to focus on network stability and not risk confounds with complicated behaviour in unrelated interactions. We are only aware of three experiments examining pairwise stability in a pure network formation game: Pantz (2006), Carrillo and Gaduh (2012), Kirchsteiger et al. (2013). Our experiment differs from these in two fundamental ways: first of all, subjects in our experiment are paid for all intermediate steps rather than only for the final network; secondly, our subjects interact in a far less structured manner.

Clearly paying all intermediate steps is necessary to test our new theory, but in addition we claim that this is also an appropriate environment in which to test other stability concepts. In the design of an experiment there is always a question of which theoretical assumptions to control and which to leave to the behaviour of subjects. Kirchsteiger et al. (2013) chose to pay only the final network, thus imposing the preferences on their subjects that were assumed by the farsighted stability concept they were interested in and making their experiment as close as possible to the theory. By paying intermediate steps, we examine whether caring only about a final network is a behavioural feature exhibited by actual subjects.8

A further reason for testing the existing concepts in an experiment which pays for intermediate steps is that, as argued above, many real-life network-formation environments involve regular payoffs, all of which should matter to the agents. For a stability concept to be useful it should apply not only to situations where all assumptions are strictly satisfied, but also in more realistic environments.

Our experiment differs from most other network experiments in allowing for free timing of moves, in the sense that players can propose, accept, and delete links at any point in the game, with incremental payoffs changing as soon as the network structure changes.9 Moreover, the process of link formation, which involves a link proposal by one party, observation of the proposal by the other party and either acceptance or rejection of the proposal, implies the possibility of communication between pairs of players.

---

8To clarify this point, consider the fact that in this experiment we also claim to test PWS which assumes myopic behaviour. It would be unusual to impose myopic preferences on subjects, i.e. paying them only for the first step of link-formation.

9Berninghaus et al. (2006) also implement their experiment with free-timing of moves and a flow of payoffs, but with unilateral link formation. Callander and Plott (2005) have treatments with free timing of moves, with the network in place after two minutes used for payment, however their motivation was to overcome coordination problems that arise from the non-cooperative game they implement.
This seems to us to not only increase realism compared to the random structure imposed in earlier experiments, but also be more in the spirit of the pairwise cooperative nature of the theory. A further advantage of our design is that stable networks can arise very quickly, with the brevity of each game allowing for a large number of repetitions and thus greater learning.\footnote{Our subjects repeat the game 20 times, compared with three or four times in other papers.}

We find that the most stable networks are the pairwise stable ones (PWS). We also find evidence of stability for a network which is not PWS, but is part of sets identified by the pairwise farsightedly stable set (PWFS), von Neumann-Morgenstern pairwise farsightedly stable set (vN-MFS), largest pairwise consistent set (LPWC), level-K farsightedly stable set (Level-K), and our one. We also find an empirically stable network which is identified only by LPWC and ours. However, LPWC (along with PWFS and Level-K) also identifies as potentially stable networks which are not observed to be so. Thus, our concept is the only one which precisely identifies the set of empirically stable networks.

The paper is organized as follows. In Section 2 we introduce our new theory of network formation. We first define the notions of path payoffs, improving and surely improving paths and introduce the concept of the cautious path stable set of networks. Then we state its existence and conditions for uniqueness. We demonstrate and compare the predictions of our new stability concept and a range of other myopic and farsighted pairwise stability concepts in two games, which are later used in our experiment. At last, we study the relationship between our new concept and other concepts of farsighted stability in the special case when players are only interested in the end-of-path payoffs. Then in Sections 3 and 4, we describe our laboratory experiment. Finally, in Section 5 we conclude. Proofs, tables and experiment instructions are provided in the Appendix.

\section{Theory}

\subsection{Networks}

Consider a network $g$ on $n$ nodes. Nodes of the network are players and links indicate bilateral relationships between players. The relationships are symmetric, or reciprocal, which means that links do not have an associated direction, and the networks is, therefore, \textit{undirected}. We say that $ij \in g$ if players $i$ and $j$ are linked in the network $g$. In the \textit{complete} network all players are linked with each other, that is, $ij \in g$ for any pair of players $ij$. In the \textit{empty} network, no pair of players is linked, that is, $ij \notin g$ for any pair of players $ij$. 

\footnote{Our subjects repeat the game 20 times, compared with three or four times in other papers.}
The set of all possible networks on $n$ nodes is denoted by $\mathbb{G}$. The network obtained by adding link $ij$ to an existing network $g$ is denoted by $g + ij$, and similarly, the network obtained by deleting link $ij$ from an existing network $g$ is denoted by $g - ij$.

### 2.2 Paths

A path from a network $g$ to a network $g'$ is a finite sequence of networks $P = \{g_1, \ldots, g_K\}$, where $g_1 = g$, $g_K = g'$ and for any $1 \leq k \leq K - 1$ either (i) $g_{k+1} = g_k - ij$ for some $ij$, or (ii) $g_{k+1} = g_k + ij$ for some $ij$, or (iii) $g_{k+1} = g_k$. We will sometimes say that path $P$ leads from $g$ to $g'$, and if $g' \in G \subseteq \mathbb{G}$, then path $P$ leads to $G$. The length of path $P$ is the number of networks in the sequence; it is denoted by $|P|$. In the definition of path $P$ here $|P| = K$.

A special path is a constant path that consists of a certain number of repetitions of the same network. A constant path that consists of $m$ repetitions of network $g$ is denoted by $g^m$. Clearly, $|g^m| = m$.

For any two paths $P = \{g_1, \ldots, g_K\}$ and $P' = \{g'_1, \ldots, g'_K\}$, we define a path $P \cup P'$ as a path that is obtained by concatenation of paths $P, P'$ in the specified order: $P'$ after $P$. That is, $P \cup P' = \{g_1, \ldots, g_K, g'_1, \ldots, g'_K\}$. Note that in general, $P \cup P' \neq P' \cup P$. But $|P \cup P'| = |P'| \cup P| = |P| + |P'|$.

Finally, for any path $P = \{g_1, \ldots, g_K\}$ and any $1 \leq k \leq K$, we define a continuation of path $P$ from position $k$ as a sequence of networks on path $P$ from network $g_k$ onward. That is, a continuation of path $P$ from position $k$ is path $P_k = \{g_k, \ldots, g_K\}$. In particular, a continuation of path $P$ from position $1$ is path $P$ itself, i.e., $P_1 = P$, and for any $k > 1$, $P = \{g_1, \ldots, g_{k-1}\} \cup P_k$.

The (infinite) set of all paths between any pair of networks in $\mathbb{G}$ is denoted by $\mathbb{P}$.

### 2.3 Path payoffs

For any player $i$, we define a path payoff as a function $\pi_i : \mathbb{P} \to \mathbb{R}$ that specifies payoff $\pi_i(P)$ that player $i$ obtains on any path $P \in \mathbb{P}$. While we do not impose any specific assumptions on the functional form of $\pi_i$, one way to think about a path payoff of player $i$ is that it is a weighted average of payoffs that player $i$ obtains in any network on the path. In that, the exact definition of the weights and of the weighted average is subject to a specific context. For example, denoting by $Y_i(g)$ a payoff that player $i$ obtains in a network $g$, a path payoff can be defined as $\pi_i(P) = Y_i(g)$ for some network $g$ on path $P = \{g_1, \ldots, g_K\}$. This definition is reasonable when player $i$ allocates positive weight to just one network on the path. In particular, if $g = g_1$, then player $i$ assigns positive weight only to the first network on the path, while if $g = g_K$, then player $i$ “cares” only about the last network. The former case is commonly assumed in
settings where players are *myopic*, such as in the definition of pairwise stability (Jackson and Wolinsky, 1996), while the latter case is suitable for the environments where players are *farsighted* and do not care about gains and losses they may incur before the final network is reached; this latter specification of payoffs is employed in the definitions of many existing farsighted stability concepts (Herings et al., 2009; Chwe, 1994). In intermediate cases, where player \( i \) is interested not only in the immediate or final payoff but also in payoffs accrued from intermediate steps, the path payoff of player \( i \) associated with path \( P \) can be defined using the exponential discounting, as 
\[
\pi_i(P) = Y_i(g_1) + \delta Y_i(g_2) + \ldots + \delta^{K-1} Y_i(g_K)
\]
for some \( \delta > 0 \), or as 
\[
\pi_i(P) = \varepsilon (Y_i(g_1) + \ldots + Y_i(g_{K-1})) + Y_i(g_K)
\]
for some \( \varepsilon > 0 \), or 
\[
\pi_i(P) = \frac{1}{K} (Y_i(g_1) + \ldots + Y_i(g_K)).
\]
This last interpretation of path payoffs as the simple arithmetic average of payoffs in all networks will be particularly useful for the discussion of the results in our experiment later.

**Example 1** Consider a set of all possible networks for the 3-player case depicted on Figure 1. These are the empty network \( g_0 \), complete network \( g_7 \), three 1-link networks \( g_1, g_2, g_3 \) and three 2-link networks \( g_4, g_5, g_6 \). The payoff of a player in each network is represented by a number next to the corresponding node.

![Figure 1: Example 1.](image)

Consider a path \( P = \{g_1, g_5, g_3\} \) that leads from one 1-link network to another 1-link network via a 2-link network. If Player 1 (Pl.1) is interested only in the final network on any path, then her path payoff associated with \( P \) is \( \pi_1(P) = Y_1(g_3) = 6 \). If, on the other hand, Player 1 weighs payoffs in all networks on a path equally, then her path payoff is the arithmetic average, 
\[
\pi_1(P) = \frac{1}{3} (Y_1(g_1) + Y_1(g_5) + Y_1(g_3)) = 20
\]
With exponential discounting, her path payoff is 
\[
\pi_1(P) = Y_1(g_1) + \delta Y_1(g_5) + \frac{\delta^2}{1-\delta} Y_1(g_3) = 30 + 24\delta +
\]
$6 \frac{\delta^2}{1-\delta}$. And if Player 1 is mostly interested in the final network but assigns a small positive weight $\varepsilon$ to intermediate networks, then $\pi_1(P) = \varepsilon (Y_1(g_1) + Y_1(g_5)) + Y_1(g_3) = 54\varepsilon + 6$. Clearly, this difference in payoff specification can lead to different predictions for network stability.

### 2.4 Improving paths

We define two special types of paths: an improving path and a surely improving path. Both of these notions will be used in the definition of our new concept of network stability that we discuss in the next section.

An improving path is a sequence of networks that can emerge when players add or sever links based on the improvement that this sequence offers relative to staying in the current network. Each network in the sequence differs from the previous by one link. If a link is added, then the two players involved must both prefer the path payoff associated with the remainder of the path (starting after the link was added to the current network) to the payoff associated with staying in the current network for the same number of steps. If a link is deleted, then it must be that at least one of the two players involved in the link strictly prefers the payoff associated with the remainder of the path.$^{11}$ As usual with pairwise deviations, the idea behind this definition is that adding a link requires a consent of both players involved, while deleting a link can be done unilaterally. The formal definition is as follows.

**Definition 1** A finite path $P = \{g_1, \ldots, g_K\}$ is an improving path if for any $1 \leq k \leq K - 1$ either

(i) $g_{k+1} = g_k - ij$ for some $ij$ such that $\pi_i(P_{k+1}) > \pi_i(g_k|P_{k+1}|)$ or $\pi_j(P_{k+1}) > \pi_j(g_k|P_{k+1}|)$, or

(ii) $g_{k+1} = g_k + ij$ for some $ij$ such that $\pi_i(P_{k+1}) > \pi_i(g_k|P_{k+1}|)$ and $\pi_j(P_{k+1}) \geq \pi_j(g_k|P_{k+1}|)$.

For a given network $g$, let us denote by $P^I(g)$ the set of all improving paths from network $g$. One useful observation is that if $P$ is an improving path from $g_1$ to $g_K$, then a continuation of $P$ from any step $k$, $1 < k \leq K - 1$, is an improving path from $g_k$ to $g_K$. That is, if $P \in P^I(g_1)$, then $P_k \in P^I(g_k)$ for any $1 < k \leq K - 1$.

Furthermore, note that for the appropriately chosen specification of path payoffs, the definition of the improving path is equivalent to the definition of the myopic improving path or farsighted improving path introduced in Jackson and Watts (2002a) and Herings et al. (2009). Indeed, if players care

---

$^{11}$Similarly, on the farsighted improving path defined by Herings et al. (2009) players compare the payoff in the final network of the path with the payoff in the current network.
only about their immediate payoff, that they obtain straight after adding or deleting a link, then 
\( \pi_i(P_{k+1}) = Y_i(g_{k+1}) \) and \( \pi_i(g_k^{[P_{k+1}]})) = Y_i(g_k) \). In this case an improving path is, in fact, the myopic improving path of Jackson and Watts (2002a). If, on the other hand, players care only about their payoff in the final network of a path, then \( \pi_i(P_{k+1}) = Y_i(g_K) \) and \( \pi_i(g_k^{[P_{k+1}]})) = Y_i(g_k) \). In this case, an improving path is the farsighted improving path of Herings et al. (2009).

**Example 2** Consider again the set of all possible networks for 3 players, and let the payoffs in each network be as shown on Figure 2. Suppose that players’ path payoff is the simple arithmetic average of their payoffs in all networks on the path. Then it is easy to see that as 30 is the absolute maximum of what the players can gain in any network, there are no improving paths starting at any of the 1-link networks: a player with payoff 30 does best for herself by simply staying in the same network rather than by following some path. On the other hand, from the empty network \( g_0 \) there exist an obvious improving path to each of the 1-link networks but there is no improving path anywhere else as there are no improving paths from 1-link networks. From each of the 2-link networks there are improving paths to two 1-link networks and nowhere else: from \( g_4 \) there are improving paths to \( g_1 \) and \( g_2 \), from \( g_5 \) to \( g_1 \) and \( g_3 \), and from \( g_6 \) to \( g_2 \) and \( g_3 \).\(^{12}\) Finally, from the complete network \( g_7 \) there exists at least one improving path to any other network, apart from the empty network. For example, \( P_1 = \{g_7,g_4,g_1\} \),

\(^{12}\)With equal weighting of all networks on the path, there is no improving path to the third 1-link network in each case, neither via another 1-link network nor via the complete network. However, if players assigned sufficiently higher weight to the final network on a path, there would exist improving paths from a 2-link network to all 1-link networks. For example, \( \{g_4,g_7,g_6,g_3\} \) would be an improving path as soon as the last network on the path was relatively more important for players than the other networks.
$P_2 = \{g_7, g_4, g_2\}$, $P_3 = \{g_7, g_6, g_3\}$ are improving paths to each of the 1-link networks, and $P_4 = \{g_7, g_4\}$, $P_5 = \{g_7, g_5\}$, $P_6 = \{g_7, g_6\}$ are improving paths to each of the 2-link networks.

Note that path $P_1$ is improving, as its continuation from step 2 strictly improves the average payoff of Player 2 ($22 < \frac{1}{2}(24 + 30)$) and the continuation from step 3 improves the average payoff of Player 1 ($18 < 30$). The payoff of Player 3 declines. Therefore, on this path Player 2 deletes the first link and Player 1 deletes the second. Moreover, due to the symmetry of players’ payoffs, Player 1 in network $g_4$ at the second step of the path is actually indifferent between deleting either of her two links. If she deletes the other link instead, then from the perspective of Player 2, it is not worth deleting the first link as it eventually reduces her average payoff ($\frac{1}{5}(24 + 6) < 22$). This implies that if no commitment can be made by Player 1 to delete the link with Player 3 and not with Player 2, then Player 2 may prefer to avoid the risk and not delete any link in the first place. These considerations are taken into account in the definition of the surely improving path that we consider next.

Example 2 hints that when full-communication and/or commitment are not possible, cautious players may abstain from deleting or adding a link on an improving path. We incorporate this idea of cautiousness in the definition of the improving path by assuming that players delete or add a link only if their payoff improves not just on this but on any improving path that follows after that. Moreover, among these paths, players take into account only those improving paths that lead to a network in set $G$, where $G$ is regarded as a stable or absorbing set. The definition of a stable set is provided in the next section. For now, it just introduces the idea of a credible threat, in the sense that players’ moves on a surely improving path can only be deterred by those of the subsequent improving paths that lead to a stable set.

To be more precise, we call an improving path surely improving relative to set $G$ if (i) whenever a link is deleted, at least one of the two players involved in the link prefers any improving path that starts after the link is deleted and leads to a network in $G$ to staying in the current network for the same number of steps, and (ii) whenever a link is added, both involved players prefer any improving path that starts after the link is added and leads to a network in $G$ to staying in the current network, with at least one of the two preferences being strict.

**Definition 2** A finite path $P = \{g_1, ..., g_K\}$ is surely improving relative to $G$ if it is an improving path and for any $1 \leq k \leq K - 1$ either\(^\text{13}\)

\(^{13}\)If $\bar{P}$ such that $\bar{P} \in P^i(g_{k+1})$ and leads to $G$ does not exist, that is, there is no credible threat that deleting or adding
(i) \( g_{k+1} = g_k - ij \) for some \( ij \) such that \( \pi_i(\tilde{P}) > \pi_i(g_k^{\tilde{P}}) \) for any \( \tilde{P} \in P^I(g_{k+1}) \) leading to \( G \) or \( \pi_j(\tilde{P}) > \pi_j(g_k^{\tilde{P}}) \) for any \( \tilde{P} \in P^I(g_{k+1}) \) leading to \( G \), or

(ii) \( g_{k+1} = g_k + ij \) for some \( ij \) such that \( \pi_i(\tilde{P}) \geq \pi_i(g_k^{\tilde{P}}) \) and \( \pi_j(\tilde{P}) \geq \pi_j(g_k^{\tilde{P}}) \), with at least one inequality being strict, for any \( \tilde{P} \in P^I(g_{k+1}) \) leading to \( G \).

For a given network \( g \), we denote by \( P^{SI}(g,G) \) the set of all paths starting at network \( g \) that are surely improving relative to \( G \). By definition, \( P^{SI}(g,G) \subseteq P^I(g) \) for any \( G \subseteq G \).

The definition of a surely improving path assumes players’ cautiousness in two respects. First, just as with max-min preferences, a decision of a player to add or delete a link is discouraged by the existence of at least one improving path starting after the player’s move and leading to \( G \) on which this player’s payoff is worse than the payoff associated with staying in the status quo network.\(^{14}\) Second, among all paths that might be followed after the link is added or deleted, players give consideration to all improving paths leading to \( G \), and not only to the surely improving ones. The latter is reasonable when players, for example, do not know how cautious or sophisticated the others are, and being cautious themselves, take into account all possibilities.

Note that such “extreme cautiousness” in players’ behaviour makes the existence of surely improving paths between networks harder than under alternative, less cautious approaches, where players consider not all but only surely improving paths or take into account the weighted average of possible improving paths. As a result, the set of networks at which no or few surely improving paths begin is larger, and this eventually leads to the existence and non-emptiness of a stable set of networks in our setting. It also implies that our stable set is larger, and networks which are not stable according to our definition cannot be stable according to these other, less cautious approaches.

Furthermore, note that players who add or delete a link on a surely improving path take into account not just the improving paths that start immediately after this link change but also all improving paths leading to \( G \) that start at any later step on the path. In particular, players who initiate the move on a surely improving path take into consideration all improving paths that start at the last network of the path and lead to \( G \), i.e., all possible improving continuations (leading to \( G \)) of the given surely improving path. Indeed, suppose that path \( P = \{g_1, ..., g_K\} \) is surely improving relative to \( G \), that

\(^{14}\)The same approach in evaluating the possibilities is exhibited by players located on a network inside (but not outside) the pairwise farsightedly stable set of networks in Herings et al. (2009). More detailed comparison is provided in Sections 2.5 and 2.8.
is, \( P \in P^{SI}(g_1, G) \). Consider that for any \( 1 < k \leq K \) and any improving path \( \tilde{P} \) starting at \( g_k \) and leading to \( G \), a path \( \{g_{k-1}\} \cup \tilde{P} \) is also an improving path leading to \( G \) but starting at \( g_{k-1} \), i.e.,
\[
\{g_{k-1}\} \cup \tilde{P} \in P^I(g_{k-1}).
\]
Then by induction, \( \{g_{k-2}, g_{k-1}\} \cup \tilde{P} \in P^I(g_{k-2}) \) and leads to \( G \) and so on. So, in general, \( \{g_l, \ldots, g_{k-1}\} \cup \tilde{P} \in P^I(g_l) \) and leads to \( G \) for any \( 1 \leq l < k - 1 \). This means that a player(s) who deletes or adds a link on the transition from \( g_{l-1} \) to \( g_l \) of a surely improving path \( P \), is guaranteed to become better off on any improving path leading to \( G \) that starts not just at \( g_l \) but also at any future network on the path.

Just as the definition of an improving path implies that any continuation of an improving path is also an improving path, the definition of a surely improving path suggests that any continuation of a surely improving path is a surely improving path, too. That is, for any path \( P = \{g_1, \ldots, g_K\} \) such that \( P \in P^{SI}(g_1, G) \), a continuation \( P_k = \{g_k, \ldots, g_K\} \) for any \( 1 < k \leq K - 1 \) is such that \( P_k \in P^{SI}(g_k, G) \). Moreover, if a path is surely improving relative to \( G \), then it is also surely improving relative to any subset of \( G \). That is, if \( P \in P^{SI}(g, G) \), then \( P \in P^{SI}(g, G') \) for any \( G' \subseteq G \), so that \( P^{SI}(g, G) \subseteq P^{SI}(g, G') \).

A slightly less straightforward pair of properties is stated by Lemma 1 and Lemma 2. The first property establishes the “transitivity” of surely improving paths, in the sense that a union of two surely improving paths, where the end of the first path is the beginning of the second, is a surely improving path. More formally, if the first path is surely improving relative to set \( G \) and the second is surely improving relative to set \( G' \) but leads to a network in \( G \), then the union of the two paths is surely improving relative to the intersection of \( G \) and \( G' \), and in fact, relative to any subset in the intersection. In particular, a union of two surely improving paths relative to the same set \( G \), where the second path leads to \( G \), is surely improving relative to \( G \) and any smaller set. In a similar way, the second property establishes that a union of two improving paths, where only the first is surely improving, is an improving path.\(^{15}\)

**Lemma 1** Suppose that \( P \in P^{SI}(g, G) \) and \( P \) leads to \( g' \), \( P' \in P^{SI}(g', G') \) and \( P' \) leads to \( g'' \in G \), and \( G \cap G' \neq \emptyset \). Then \( P'' = P \cup P'_2 \in P^{SI}(g, G'') \) for any \( G'' \subseteq G \cap G' \).

**Lemma 2** If \( P \in P^{SI}(g, G) \) and \( P \) leads to \( g' \), and \( P' \in P^I(g') \) and \( P' \) leads to \( g'' \in G \), then \( P'' = P \cup P'_2 \in P^I(g) \).

\(^{15}\)The proof of Lemma 2 is a subproof of Lemma 1 and is, therefore, omitted. Indeed, in order to show that \( P'' \) is a surely improving path in Lemma 1, one needs to verify, in particular, that it is an improving path, and this part of the proof only requires that the first of the two improving paths is surely improving. The details are available from the authors upon request.
To demonstrate the notion of a surely improving path, consider again the 3-player case of Example 2. In this example, all improving paths that start at the empty network and at each of the 2-link networks are at the same time surely improving relative to any set, as no threat exists that a player(s) who adds or deletes a link on these paths will become worse off. On the other hand, all improving paths that start at the complete network are not surely improving relative to $G$ as soon as $G$ contains all 1-link networks. The reason for this is described in Example 2. Namely, any player deleting a link at the first step cannot be sure that an improving path which will be followed after that will make her better off.

In Section 2.7 we will show that the existence of an improving but not surely improving path from the complete to a 1-link network leads to the complete network being unstable according to many existing farsighted stability concepts (PWFS, vN-MFS and Level-K) but stable according to our new concept.

2.5 Cautious path stable sets of networks

We now introduce a new concept of network stability, that we will call the cautious path stable set, or briefly, the CPS set. The definition of the cautious path stable set $G$ requires that it is the minimal set that satisfies the internal and external stability in the sense that (1) there does not exist a surely improving path relative to $G$ between any pair of networks in the set, and (2) there exists a surely improving path relative to $G$ from any network outside the set leading to some network in the set. Formally, the cautious path stable set of networks is defined as follows.

**Definition 3** A set of networks $G \in \mathbb{G}$ is cautious path stable (CPS) if

1. \( \forall g \in G \ \not\exists P \in P^{SI}(g,G) \) such that $P$ leads to $G \setminus \{g\}$;
2. \( \forall g' \in \mathbb{G} \setminus G \ \exists P \in P^{SI}(g',G) \) such that $P$ leads to $G$;
3. \( \forall G' \subset G \) at least one of conditions (1), (2) is violated by $G'$.

Conditions (1) and (2) of the definition imply that the networks within a stable set are robust to perturbations, and condition (2) also means that any cautious path stable set is not empty. Moreover, condition (2) together with the transitivity of surely improving paths (see Lemma 1) implies that if $G$ is a cautious path stable set, then any surely improving path relative to $G$ starting at a network in $G$ must be such that it eventually leads back to exactly the same network in $G$. Indeed, by condition (1) there are no surely improving paths relative to $G$ between any two networks in $G$, and any surely improving path that leads from network $g \in G$ to a network outside $G$ has a continuation – by condition (2) – back to set $G$. This continuation leads back to exactly the same network $g$, as if it lead to any other

---

16Recall that there are no improving paths that start at a 1-link network.
network in \( G \), we would have a contradiction to condition (1). Therefore, a cautious path stable set represents an “absorbing” set of networks, in the sense that once entered (by a surely improving path), it cannot be left without coming back to exactly the same network.

Our definition of the cautious path stable set is similar to the definition of the von Neumann-Morgenstern pairwise farsightedly stable set (vN-MFS set) introduced in Herings et al. (2009).\(^{17}\) Just as our concept, vN-MFS imposes internal and external stability, and no proper subset of the stable set satisfies these two conditions, but instead of surely improving paths the definition of vN-MFS set uses the notion of the so-called farsighted improving paths (Herings et al., 2009). As a result, two important differences are in order. First, for a generic definition of path payoffs, our concept assumes that players are interested not only in their payoffs in the final network of the improving path but also in their payoffs on intermediate steps. Second, players in our setting are assumed to be cautious and add or delete links only if their path payoffs are certain to increase on any improving path leading to \( G \) that might be followed after the link change and not just on the given improving path. Moreover, while vN-MFS sets do not always exist, the existence of the cautious path stable set is easy to prove.\(^{18}\)

In Section 2.8 we will show that this new definition of a stable set, in a special case when players care only about their final network payoffs, is also similar in spirit to the definition of the pairwise farsightedly stable set (PWFS set) of Herings et al. (2009). Essentially the same three conditions – deterrence of external deviations, external stability and minimality – are satisfied by our stable set, but the important difference is that the external stability in our definition requires the existence of not just an improving but surely improving path relative to \( G \) from any network outside \( G \) to a network in \( G \). Therefore, players are cautious and consider the consequences of adding and deleting a link not only when they are in a network inside \( G \) but also when they are outside \( G \). In a sense, by introducing the requirement that a path from a network outside \( G \) to a network in \( G \) must be surely improving, we “add more cautiousness” to players’ behaviour relative to what is assumed in Herings et al. (2009).

More generally, the key features underlying the concept of the cautious path stable set – players’ cautiousness and generic definition of path payoffs – distinguish this concept from other concepts of farsighted pairwise stability. How these differences matter for the predictions of our stability concept relative to those of other concepts will be demonstrated on the example of two network formation games

\(^{17}\)vN-MFS is based on the original definition of the von Neumann-Morgenstern stable set (von Neumann and Morgenstern, 1944).

\(^{18}\)See Proposition 2 in the next section.
in Section 2.7.

Note that if set $G$ consists of a single network, then conditions (1) and (3) of the definition of a cautious path stable set are trivially satisfied. As a result, stability of $G$ is fully determined by condition (2).

**Proposition 1** The set $\{g\}$ is cautious path stable if and only if $\forall g' \neq g \exists P \in P_{SI}(g', \{g\})$ such that $P$ leads to $g$.

Furthermore, the minimality of a cautious path stable set suggests that if $\{g\}$ is a cautious path stable set, then it does not belong to any other stable set. But there may exist other cautious path stable sets that do not contain $g$. In the next section, we will consider this question more broadly. Namely, we will first discuss the existence of a cautious path stable set and then provide some easy to verify conditions for the set to be cautious path stable and the unique cautious path stable set.

### 2.6 Existence of CPS set. Characterization of CPS set

The first important result establishes the existence of a cautious path stable set in any (pairwise cooperative) network formation game.

**Proposition 2** A cautious path stable set of networks exists.

The proof of Proposition 2 proposes an algorithm to construct one such set. First, we observe that the whole network space $\mathbb{G}$ trivially satisfies condition (2) of the definition of a cautious path stable set. If it also satisfies condition (1), then the proof is completed, as either $\mathbb{G}$ or one of its proper subsets is the minimal set that satisfies conditions (1) and (2). If condition (1) is not satisfied, then there must exist at least one network $g \in \mathbb{G}$ and a surely improving path relative to $\mathbb{G}$ starting at $g$ such that this path leads to some other network in $\mathbb{G}$. Then we consider a set $G_1 = \mathbb{G} \setminus \{g\}$, which again, trivially satisfies condition (2) but might not satisfy condition (1). If condition (1) is not satisfied, then we consider the next, smaller set $G_2 = G_1 \setminus \{g_1\}$, where $g_1$ is a network from which a surely improving path relative to $G_1$ leads to another network in $G_1$. We show that set $G_2$ satisfies condition (2), as for any network outside $G_2$, there exists a surely improving path relative to $G_2$ that leads to $G_2$, either via network $g_1$ (by Lemma 1) or “directly”. If condition (1) is also satisfied, then the proof is completed; otherwise, we reduce the set even further, by “taking out” network $g_2$ from which there exists a surely improving path relative to $G_2$ to another network in $G_2$, etc. This procedure generates a decreasing sequence of
proper subsets, where each subset satisfies condition (2). The existence of a cautious path stable set is then established by observing that a) the last element of this sequence, that satisfies both conditions, (1) and (2), is certain to exist because the cardinality of \( \mathcal{G} \) is finite, and a set consisting of a single network trivially satisfies condition (1), and b) either this set itself or one of its proper subsets is the minimal set that satisfies (1) and (2), so that condition (3) also holds.

The algorithm provided in the proof of Proposition 2 constructs one cautious path stable set. But the outcome of this algorithm, in general, depends on the order in which the networks, violating condition (1), are taken out of the set and on the last-step choice of the minimal subset that satisfies conditions (1), (2) and (3). Therefore, a cautious path stable set might not be unique. The next proposition provides two simple conditions that are sufficient for the set to be the unique cautious path stable set.

**Proposition 3** If for every \( g \in \mathcal{G} \) \( P^I(g) = \emptyset \) or any \( P \in P^I(g) \) is such that \( P \) leads to \( g \), and for every \( g' \in \mathcal{G} \setminus \mathcal{G} \) \( \exists P \in P^{SI}(g', G) \) such that \( P \) leads to \( G \), then \( G \) is the unique cautious path stable set.

**Proof.** First, it is easy to see that \( G \) satisfies conditions (1) and (2) of the definition of the cautious path stable set. Second, no proper subset of \( G \) satisfies condition (2), therefore, \( G \) is minimal in the sense of condition (3) and hence, it is a cautious path stable set. Moreover, due to the same condition (2), \( G \) is a subset of any cautious path stable set. Then by minimality, \( G \) is the unique cautious path stable set. ■

Later on, in Section 2.7, we will show by means of examples that Proposition 3 cannot be extended to an “if and only if” statement. Our next proposition provides another couple of simple conditions that describe a cautious path stable set. These conditions are less restrictive than those required for uniqueness in Proposition 3 and they are simpler than those in the definition of the cautious path stable set.

**Proposition 4** If for every \( g \in \mathcal{G} \) any \( P \in P^I(g) \) is such that \( P \) leads to \( \{g\} \) or to \( \mathcal{G} \setminus \mathcal{G} \), and for every \( g' \in \mathcal{G} \setminus \mathcal{G} \) \( \exists P \in P^{SI}(g', G) \) such that \( P \) leads to \( G \), then \( G \) is a cautious path stable set.

**Proof.** First, observe that conditions (1) and (2) are trivially satisfied. Moreover, condition (3) is satisfied, too, as for any proper subset \( G' \subsetneq G \), condition (2) does not hold: \( \forall g' \in \mathcal{G} \setminus G' \exists P \in P^{SI}(g', G') \) such that \( P \) leads to \( G' \), as there does not exist even a simple improving path from \( g' \) to \( G' \). ■
2.7 Examples of CPS sets. Game 1 and Game 2

In this section we show that in the network formation games of Example 1 and 2, there exists a unique cautious path stable set of networks. We also demonstrate how the predictions of this stability concept differ from the predictions of other concepts of farsighted and myopic pairwise stability. These predictions for the two games (Game 1 and Game 2) will then be tested empirically, in the experiment that we discuss in Sections 3 and 4.

Game 1 (G1) Consider first the network formation game of Example 1, where players’ payoffs in every network are as shown on Figure 3. In our experiment we will refer to this game as G1.\footnote{In both games considered in this section, network payoff allocation across players is anonymous, that is, payoffs depend only on player positions in the network, and not player label. This is consistent with a standard approach in the literature (Dutta et al., 2005; Herings et al., 2009) and goes back to the original proposal of anonymity by Jackson and Wolinsky (1996) as a basic property of payoff allocation rules in networks.} As before, suppose that players’ path payoffs are the arithmetic average of their payoffs in all networks on a path. Below we will show that the unique cautious path stable set of networks is $G = \{g_1, g_2, g_3, g_7\}$.

First, we observe that the complete network, $g_7$, must belong to any cautious path stable set, as there are no improving paths from $g_7$ to any other network. Appendix A provides the details of the argument. Then, since $g_7$ belongs to each stable set, all 1-link networks must belong to each stable set, too, as no path starting at a 1-link network is surely improving relative to a set containing $g_7$. Indeed, consider that any improving path from a 1-link network involves either deleting or adding a link at the first step. If the link is deleted, then for such a path to be improving it must at some point leave the set of empty and 1-link networks because 30 is the maximal payoff of a player in this set. Hence, a 2-link

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{game1.png}
\caption{Game 1.}
\end{figure}
network is formed from a 1-link network at some step of this path. However, such step cannot belong to a surely improving path relative to a set containing \( g_7 \), because from a 2-link network, there exists a 1-step improving path to \( g_7 \), on which a player with payoff 30 in the 1-link network, who added a link, becomes worse off. Indeed, on the path from the 2-link network to the complete network, the average payoff of this player is \( \frac{1}{2}(32 + 22) = 27 \), which is smaller than her payoff associated with staying in the 1-link network for the same two steps. Hence, no path that involves deleting a link from a 1-link network at the first step is surely improving. By the same logic, a path that involves adding a link to a 1-link network at the first step is not surely improving either, as after such first step, a player with the initial payoff of 30 may become worse off.

Now, given that the complete and all 1-link networks belong to any cautious path stable set, all other networks are unstable, as there exists a surely improving path relative to this set leading from these networks either to the complete or to a 1-link network. Clearly, a one-step path from the empty to any 1-link network is surely improving and so is a one-step path from a 2-link to the complete network. The latter follows from the fact that a step from a 2-link network to the complete network is immediately beneficial for both players adding a link, and no further improving paths start at the complete network. Thus, \( G = \{ g_1, g_2, g_3, g_7 \} \) is a cautious path stable set, and by construction, it is unique.

The predictions of other farsighted and myopic stability concepts in Game 1 are either the same (vN-MFS and LPWC) or indicate, in addition, the potential stability of 2-link networks (PWFS), or identify just one, complete network as stable (PWS, PWMS and Level-K, for all \( K \geq 1 \)). The fact that in addition to set \( G \), the concept of PWFS identifies several other stable sets that include 2-link networks is a result of certain incautiousness or optimism assumed on the part of the players. For example, the set \( G' = \{ g_1, g_6, g_7 \} \) is PWFS because of the existence of a farsighted improving path (in terminology of Herings et al., 2009) from 1-link networks \( g_2 \) and \( g_3 \) to \( g_6 \).\(^{20}\) However, the fact that Player 3 in \( g_2 \) and \( g_3 \) is willing to add a link on this path assumes that she disregards the possibility that in \( g_6 \), the unconnected players have an incentive to add the last missing link, which would decrease her payoff. Using our definitions, this particular farsighted improving path is improving but not surely improving.

**Game 2 (G2)** Now, consider a network formation game of Example 2, where payoffs in every network are as shown on Figure 4. In the experiment we will refer to this game as G2. If as before players’ path

\(^{20}\text{One can show that } g_6 \text{ is the only network in } G' \text{ that can be reached from } g_2 \text{ and } g_3 \text{ via a farsighted improving path.} \)}}
payoff is the arithmetic average of their payoffs in all networks on a path, then the unique cautious path stable set of networks is $G = \{g_1, g_2, g_3, g_7\}$, the same as in Game 1. Indeed, from the discussion in Example 1 it follows that all 1-link networks must belong to any stable set, as there are no improving paths starting at these networks. And as soon as all 1-link networks belong to a stable set, the complete network must belong to each stable set, too, since no path starting at the complete network is surely improving relative to a set containing all 1-link networks.  

On the other hand, the empty network and all 2-link networks are such that there exists a surely improving path relative to $G$ from each of them to a 1-link network. Then Definition 3 immediately implies that $G = \{g_1, g_2, g_3, g_7\}$ is a cautious path stable set and this set is unique.

Other farsighted and myopic stability concepts, namely, PWS, PWMS, PWFS, vN-MFS, LPWC and Level-K (for all $K \geq 1$), also identify each of the 1-link networks as stable but none of them, apart from LPWC, identifies the complete network as stable. The predictions of LPWC, instead, turn out to be very broad: all but the empty network belong to LPWC set, so that even the 2-link networks are identified as stable. The reason why the complete network is not stable according to most farsighted stability concepts has to do with the fact that there exists a farsighted improving path (a combination

---

21 Recall that any player initiating a move from the complete network cannot be sure that an improving path from a two-link network will not lead to a “bad” 1-link network, which would make this player worse off.

22 The reason why the 2-link networks are identified as stable by LPWC set but not by our concept has to do with the fact that payoffs in intermediate networks on a path matter for players in our setting but not in the setting of Herings et al. (2009). Even if players are cautious in both settings, – when only the final network payoffs are important and when intermediate payoffs matter, too, – a certain path that can be taken after a player in the 2-link network deletes a link and that reduces this player’s payoff, is considered to be farsighted improving in the former setting but not in the latter. The details are available from the authors.
of farsighted improving paths of length at most $K$), as defined in Herings et al. (2009, 2014), from the complete network to each of the 1-link networks. That, according to Herings et al. (2009, 2014), means that players in the complete network have an incentive to delete a link in order to reach a higher payoff in one of the 1-link networks within a stable set. In our setting, improving paths from the complete to the 1-link networks also exist but none of them is surely improving. Therefore, cautious players do not risk deleting a link in the complete network. As regards the myopic stability concepts, PWS and PWMS, they do not identify the complete network as stable, because deleting a link by either player increases her immediate payoff.

The predictions of different farsighted and myopic stability concepts in Games 1 and 2 are summarized below:

<table>
<thead>
<tr>
<th>Concept</th>
<th>Game 1</th>
<th>Game 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>PWS</td>
<td>$g_7$</td>
<td>$g_1, g_2, g_3$</td>
</tr>
<tr>
<td>PWMS set</td>
<td>${g_7}$</td>
<td>${g_1, g_2, g_3}$</td>
</tr>
<tr>
<td>PWFS set</td>
<td>${g_1, g_2, g_3, g_7}$, ${g_1, g_6, g_7}$, ${g_2, g_5, g_7}$, ${g_3, g_4, g_7}$, ${g_4, g_5, g_7}$, ${g_1, g_6, g_7}$, ${g_5, g_6, g_7}$</td>
<td>${g_1}, {g_2}, {g_3}$</td>
</tr>
<tr>
<td>vN-MFS set</td>
<td>${g_1, g_2, g_3, g_7}$</td>
<td>${g_1}, {g_2}, {g_3}$</td>
</tr>
<tr>
<td>LPWC set</td>
<td>${g_1, g_2, g_3, g_7}$</td>
<td>${g_1, g_2, g_3, g_4, g_5, g_6, g_7}$</td>
</tr>
<tr>
<td>Level-K stable set</td>
<td>$K \geq 1: {g_7}$</td>
<td>$K = 1: {g_1, g_2, g_3}$</td>
</tr>
<tr>
<td>CPS set$^{23}$</td>
<td>${g_1, g_2, g_3, g_7}$</td>
<td>${g_1, g_2, g_3, g_7}$</td>
</tr>
</tbody>
</table>

2.8 Special case: Only the final network payoffs matter

Let us consider the special case, in which the path payoff function of each player is defined as $\pi_i(P) = Y_i(g_K)$, where $g_K$ is the final network of path $P$, and $Y_i(g_K)$ is the payoff of player $i$ in this network. With such payoff specification, players only care about the payoffs that they obtain in the last network of a path and ignore the gains and losses that they incur before the last network is reached. The reason why we are interested in this particular case, is that it allows us to establish some general regularities in the relationship between cautious path stable sets and sets identified as stable by other farsighted stability concepts, which adopt exactly the same, end-of-path payoff specification.

To begin with, note that our definitions of improving and surely improving path can be simplified as for any path $P$ and any step $1 \leq k \leq K - 1$ on the path, $\pi_i(P_{k+1}) = Y_i(g_K)$ and $\pi_i(g^{|P_{k+1}|}) = Y_i(g_k)$. In
fact, with these payoffs, the definition of the improving path becomes identical to the definition of the farsighted improving path in Herings et al. (2009). For convenience, in the following we will denote by $F^I(g)$ the “ends” of all improving paths that start at network $g$, that is, the set of all networks that can be reached from $g$ via an improving path. Similarly, by $F^{SI}(g, G)$ we will denote the set of all networks that can be reached from network $g$ via a path that is surely improving relative to $G$. By analogy with paths, the set of networks that can be reached from $g$ via a surely improving path is a subset of all networks that can be reached via an improving path, i.e., $F^{SI}(g, G) \subseteq F^I(g)$ for any $G \subseteq G$.

Furthermore, rephrasing Lemmas 1 and 2 about the properties of paths obtained by concatenation of surely improving and improving paths, in the setting where only the final network payoffs matter, we obtain that 1) if $g' \in F^{SI}(g, G)$ and $g'' \in F^{SI}(g', G') \cap G$, where $G \cap G' \neq \emptyset$, then $g'' \in F^{SI}(g, G'')$ for any $G'' \subseteq G \cap G'$, and 2) if $g' \in F^{SI}(g, G)$ and $g'' \in F^I(g') \cap G$, then $g'' \in F^I(g)$.

Using this new notation, we can also rewrite the definition of a cautious path stable set of networks. We will refer to it as a cautious final-network stable set, or briefly, the CFNS set, so as to distinguish this definition from the generic Definition 3 in Section 2.5.

**Definition 4** A set of networks $G \subseteq G$ is cautious final-network stable (CFNS) if (1) $\forall g \in G \ F^{SI}(g, G) \cap G = \emptyset$; (2) $\forall g' \in G \setminus G \ F^{SI}(g', G) \cap G \neq \emptyset$; (3) $\forall G' \subseteq G$ at least one of conditions (1), (2) is violated by $G'$.

Note that the first condition uses the fact that $F^{SI}(g) \cap \{G \setminus \{g\}\} = F^{SI}(g) \cap G$. It follows immediately from the observation that when players care only about their payoffs in the final network, no surely improving and even simple improving path can lead from a network to itself, that is, $g \notin F^I(g)$.

Clearly, all results proved for the cautious path stable set continue to hold in this special case. Most importantly, a cautious final-network stable set always exists and if for every $g \in G \ F^I(g) = \emptyset$, while for every $g' \in G \setminus G \ F^{SI}(g', G) \cap G \neq \emptyset$, then $G$ is the unique cautious final-network stable set. Moreover, Proposition 1 can now be formulated using the notion of improving rather than surely improving paths: the set $\{g\}$ is cautious final-network stable if and only if $\forall g' \neq g \ g \in F^I(g')$.\(^{24}\)

Note that Definition 4, stated in terms of network sets $F^{SI}$ rather than path sets $P^{SI}$, is even more similar to the definition of the vN-MFS set than the original formulation.\(^{25}\) However, the important difference remains. As before, our stability concept considers surely improving, and not just improving

---

\(^{24}\)Note that when $G = \{g\}$ and players care only about their final network payoffs, any improving path to $g$ is surely improving relative to $G$.

\(^{25}\)Recall that a set of networks $G \in G$ is vN-MFS if (i) $\forall g \in G \ F^I(g) \cap G = \emptyset$ and (ii) $\forall g' \in G \setminus G \ F^I(g') \cap G \neq \emptyset$. 

paths, which assumes that players are cautious and add or delete links only if their final payoff is guaranteed to improve compared to the status quo, irrespective of which improving path to the stable set is followed after the link change.

Moreover, when players care only about the final network payoffs, our stability concept turns out to have an alternative interpretation that reveals its similarity to the concept of PWFS. This alternative interpretation is obtained by requiring the deterrence of external deviations, external stability and minimality – the close counterparts of the corresponding conditions in the definition of the PWFS set.

To be more precise, a set of networks $G$ is cautious final-network stable if and only if (i) all possible pairwise deviations from any network $g \in G$ to a network outside $G$ are deterred by a credible threat of ending up worse off or equally well off, (ii) there exists a surely improving path relative to $G$ from any network outside the set leading to some network in the set, and (iii) there is no proper subset of $G$ that satisfies conditions (i) and (ii).

**Proposition 5** The set $G$ is cautious final-network stable if and only if three conditions hold:

(i) $\forall \ g \in G,$

(ia) $\forall ij \notin g$ such that $g + ij \notin G,$ $\exists g' \in F^I(g + ij) \cap G$ such that $(Y_i(g'), Y_j(g')) = (Y_i(g), Y_j(g))$

or $Y_i(g') < Y_i(g)$ or $Y_j(g') < Y_j(g),$

(ib) $\forall ij \in g$ such that $g - ij \notin G,$ $\exists g', g'' \in F^I(g - ij) \cap G$ such that $Y_i(g') \leq Y_i(g)$ and $Y_j(g'') \leq Y_j(g),$

(ii) $\forall g' \in G \ F^SI(g', G) \cap G \neq \emptyset,$

(iii) $\forall G' \subset G$ at least one of conditions (ia), (ib), (ii) is violated by $G'.$

Condition (i) of the proposition suggests that when players are in a network inside $G,$ they do not have incentives to add or delete a link which would lead to a network outside $G,$ as there exists a risk that after such a deviation some improving path will be followed that leads to $g' \in G,$ where the payoff of these players is worse than or equal to their payoff in the status quo. This means that players in a network inside $G$ are cautious, as they compare their current payoff to the (credible) worst-case scenario in case of a deviation. In exactly the same way, condition (ii) implies that players are also cautious when they are in a network outside $G.$ From any network outside $G$ there must exist a surely improving
path leading to some network in $G$, which means that players are only willing to add or delete a link on the path if after that move, their payoff is certain to increase.

This cautiousness of players’ behaviour assumed in the second, external stability condition is where the key difference from the concept of PWFS comes in. According to the corresponding condition in the definition of the PWFS set, when players are in a network outside $G$, they behave rather incautiously or optimistically, or otherwise, have the possibility of full-communication and commitment, because they rely on the *existence* of some farsighted improving path that leads to a network in $G$ (and improves their payoffs), but disregard the possibility of potentially “bad” diversions from this path. Therefore, by demanding that a path from a network outside $G$ to a network in $G$ must be *surely* improving, our concept of cautious final-network stability “adds more cautiousness” to players’ behaviour relative to what is assumed in Herings et al. (2009).

The definition of the cautious final-network stable set and Proposition 5 allow us to establish some regularities in the relationship between the cautious final-network stable sets and sets identified as stable by other pairwise farsighted stability concepts, in particular, PWFS, vN-MFS and LPWC (Herings et al., 2009). First, a simple implication of Proposition 5 is that any cautious final-network stable set includes at least one PWFS set as a subset. This follows from the fact that while both stable sets satisfy the same condition regarding the deterrence of external deviations (the first condition of Proposition 5), the cautious final-network stable set satisfies a stronger external stability condition.

**Proposition 6** For any cautious final-network stable set $G^*$, there exists a PWFS set $G$ such that $G \subseteq G^*$ and there does not exist a PWFS set $G'$ such that $G^* \subset G'$.

**Proof.** The proof is straightforward. Any cautious final-network stable set $G^*$ satisfies conditions (i) and (ii) in the definition of the PWFS set, as condition (i) is identical to the one of Proposition 5 and condition (ii) is weaker than the corresponding external stability condition of Proposition 5. If $G^*$ also satisfies the minimality condition (iii) of PWFS, then it is PWFS. Otherwise, there exists a proper subset of $G^*$ that satisfies all three conditions and hence, is PWFS. To prove the second part of the

---

26 More formally, recall that by definition of PWFS set in Herings et al. (2009), being in a network inside $G$ means that players do not have incentives to deviate to a network outside $G$, as after such a deviation, there exists a farsighted improving path that leads back to $G$ and makes these players worse off or equally well off. On the other hand, being in a network outside $G$ means that there exists some farsighted improving (but not necessarily surely improving) path that leads to $G$.

27 Indeed, if $G^*$ is not a minimal set that satisfies conditions (i) and (ii), then there must exist $G' \subset G^*$ that satisfies these two conditions. Similarly, if $G'$ is not a minimal set that satisfies (i) and (ii), then there must exists a proper subset of $G'$ that satisfies both conditions, etc. As the cardinality of set $G^*$ is finite, the sequence of thus constructed subsets of $G^*$ is finite, and the last, “smallest” subset in this sequence is minimal, that is, satisfies all three conditions.
proposition observe that the existence of a PWFS set $G'$ such that $G^* \subset G'$ would imply that $G \subset G'$, where set $G$ is also PWFS. However, this contradicts condition (iii) of minimality that any PWFS set should satisfy.

Note that Proposition 6 cannot be extended to a claim that $G^* \subset G'$ holds for any PWFS set $G$. That is, given a PWFS set, one cannot always find a cautious final-network stable set to which this PWFS set belongs. This can be demonstrated by Game 1 discussed in the previous section, where the unique cautious final-network stable set turns out to be the same as the cautious path stable set with the average path payoffs. Recall that in Game 1, many PWFS sets are not subsets of the cautious final-network stable set. Intuitively, the reason for that is suggested by Proposition 5. While the external stability condition (ii) of this proposition allows for more networks in the stable set than the corresponding condition in the definition of the PWFS set, as more networks are added to a given PWFS set to meet this condition, some other networks may become “redundant” due to the minimality condition (iii). However, if $G$ is the unique PWFS set (in which case it is also the unique vN-MFS set), then Proposition 6 suggests that $G$ must be a subset of any cautious final-network stable set.

**Corollary 1** If $G$ is the unique PWFS set (and the unique vN-MFS set), then for any cautious final-network stable set $G^*$, $G \subseteq G^*$.

Next, we observe that when a cautious final-network stable set $G$ satisfies an additional constraint, that no improving paths exist between any two networks in $G$, then $G$ is PWFS and also vN-MFS.

**Proposition 7** If $G$ is a cautious final-network stable set such that $\forall g \in G \quad F^I(g) \cap G = \emptyset$, then $G$ is a PWFS set and a vN-MFS set.

**Proof.** First, by condition (2) of the definition of the cautious final-network stable set, $\forall g' \in G \setminus G \quad F^{SI}(g', G) \cap G \neq \emptyset$. As $F^I(g') \supseteq F^{SI}(g', G)$ for any $G$, we have that $F^I(g') \cap G \neq \emptyset$. This, together with the fact that $\forall g \in G \quad F^I(g) \cap G = \emptyset$, implies that $G$ is a vN-MFS set by definition and a PWFS set by Theorem 3 of Herings et al. (2009, p. 533).

The converse of Proposition 7 is, in general, not true. That is, it is not always the case that a PWFS set or a vN-MFS is at the same time a cautious final-network stable set. For example, in Game 1, Proposition 7 applies but the converse is not true: there are seven PWFS sets and only one of them is cautious final-network stable.

---

28 Moreover, the newly added networks may not satisfy condition (i) of the PWFS, i.e., the condition that all external deviations must be deterred.

The next statement shows that if the additional constraint imposed on networks in a cautious final-network stable set is even stronger than the one in Proposition 7, then a cautious final-network stable set is the unique PWFS and vN-MFS set. This condition requires that not only there are no improving paths between networks in the set but also there are no other improving paths starting at networks in the set and leading elsewhere. Note that by Proposition 3, this condition also means that the cautious final-network stable set is itself unique.

**Proposition 8** If $G$ is a cautious final-network stable set such that $\forall g \in G \ F^I(g) = \emptyset$, then $G$ is the unique cautious final-network stable, PWFS and vN-MFS set.

**Proof.** First, by Proposition 7, $G$ is a PWFS set and vN-MFS set. Moreover, as $F^I(g) = \emptyset$, the external stability condition in the definition of all concepts (CFNS, PWFS and vN-MFS) implies that $G$ must be a subset of any stable set. Then by minimality condition inherent to any definition, $G$ is the unique cautious final-network stable, PWFS and vN-MFS set. ■

Moreover, since the internal stability condition $F^I(g) \cap G = \emptyset$ of Proposition 7 is automatically satisfied when $G$ consists of a single network, Propositions 7 and 8 lead to the following simple corollary.

**Corollary 2** If $\{g\}$ is a cautious final-network stable set, then it is also a PWFS and vN-MFS set. If in addition, $F^I(g) = \emptyset$, then $\{g\}$ is the unique cautious final-network stable, PWFS and vN-MFS set.

Finally, let us consider the relationship between the concepts of the cautious final-network stable set and the LPWC set. The latter requires that both external and internal deviations are deterred, and it also satisfies the external stability condition, identical to condition (ii) of the PWFS set. In contrast, our concept does not require internal deviations to be deterred but imposes a stricter external stability condition. As a result, a general relationship between the predictions of the two concepts is hard to derive.\(^{30}\) However, in the spirit of the analogous result in Herings et al. (2009), we can show that if a network is not in the LPWC set, then it cannot be a cautious final-network stable set.

**Proposition 9** If $\{g\}$ is a cautious final-network stable set, then $g$ belongs to the LPWC set.

The proof employs the iterative procedure to find the LPWC set, proposed by Chwe (1994). Both the procedure and the proof are provided in the Appendix.

\(^{30}\)The same concern is raised in Herings et al. (2009), who argue that the PWFS sets and LPWC sets need not be consistent.
To conclude the discussion of the special case, where players are only interested in their end-of-path payoffs, we note that our theory of cautious farsightedness can be easily modified to address the case when players have limited foresight, that is, only look a few steps ahead. This can be done by considering improving and surely improving paths of length no longer than certain $K \geq 1$, and defining a cautious final-network stable set, or more generally, a cautious path stable set in terms of these paths. Similar adjustments to the concept of PWFS are proposed by Herings et al. (2014) and Morbitzer et al. (2011), that consider level-$K$ farsighted stability and $K$-step pairwise stability. In this paper, we focused on the concepts that assume perfect foresight, and we leave the theoretical and empirical investigation of the alternative approach to future research.\footnote{One drawback of the limited farsightedness approach is that a stable network or a set of networks may not exist (see Morbitzer et al., 2011) or stable sets are “non-monotonic”, in the sense, that a certain network can be identified as stable at low levels of farsightedness, unstable at medium levels, and stable again at high levels of farsightedness (see Herings et al., 2014).}

3 Experimental Design

3.1 Experimental Games

The games we implement are precisely those discussed in Section 2.7.

As we are testing concepts of farsighted stability which require subjects to understand not only the payoff structure of the games, but also chains of others’ reactions, it is crucial that they fully understand the environment. For this reason we decided to keep things as simple as possible by using three player games, and making all networks satisfy anonymity (i.e. all players in a symmetric position receive the same payoff – in one network and also, across networks of the same type).

3.2 Hypotheses

Each theoretical stability concept divides the set of all networks into those which are stable and unstable as shown in section 2.7. Thus, rather than formally define a hypothesis for each concept, it is more straightforward to identify the networks which are stable, and compare this set of empirically stable networks to the sets that are predicted theoretically. For a concept to be validated, the empirically and theoretically stable sets should coincide precisely.\footnote{Note that this is different from testing outcomes in a multi-equilibria environment, where failure to observe one of the equilibria does not invalidate the theory. We are not asking which networks will arise, but which networks are stable given that they have been entered, which is why all networks identified by a concept must be stable.} We define a type of network as being stable if such
a network is more likely to remain in place for the next payment than not:

**Definition 5** A type of network is *stable* if, conditional upon being paid in the current period, the probability of it being paid in the next is greater than 0.5.

In our first set of statistical tests, we test the hypotheses that this probability of remaining in each type of network is less than or equal to a half, and consider a type of network to be stable if this hypothesis is rejected.

We are also interested in whether networks that are myopically stable are more stable than those that require farsighted behaviour. In a second set of tests, we compare the relative stability of each type of network, predicting that networks identified by myopic concepts are more stable than those predicted only by farsighted concepts.

### 3.3 Procedural Details

The playing screen seen by subjects is shown in the instructions in the Appendix. All subjects saw themselves represented as a green circle at the bottom of the screen and the other two players as blue circles at the top.

Links could be formed between two players in the following way. A subject could indicate they were willing to form a link with another player by clicking on the appropriate blue dot, resulting in a pink arrow pointing towards the other player. Clicking again would undo this action. If two players had both clicked on each other then a link was formed and shown in red. Links could be broken by either of the parties clicking on the other, leaving a pink arrow pointing towards the player who had broken the link. All of these actions could be taken at any time.

Payoffs were made at one second intervals for 30 seconds according to the network described by the red links. Each player’s per-second payoffs were displayed next to their circle. The total points accrued were displayed and updated throughout the round, as was the number of seconds remaining.

The game was repeated 20 times with stranger matching. The starting network was randomly determined at the beginning of a round.

Before commencing the paid rounds, subjects completed a detailed tutorial familiarizing them with the interface, and played three practice rounds with payoffs different from the game of interest.

---

33 We cannot simply use the average duration of networks as the basis of a measure of stability because the final network in every game is censored.

34 We recognise that this definition is somewhat arbitrary, but we feel it to be quite natural. All our results are robust to increasing the defining probability to 0.6.
The computerized experiments were programmed in Z-tree (Fischbacher, 2007) and took place in the Vienna Center for Experimental Economics. For each game, four sessions were conducted, each consisting of 18 subjects divided into two matching groups, giving us a total of 16 independent observations. One randomly chosen period was paid, with every 45 points exchanged for 1 Euro. Sessions lasted approximately one hour.

4 Results

We begin this section by giving a descriptive overview of the data before proceeding to formal statistical tests. The first set of tests ask whether or not each type of network is “stable” using the absolute definition of stability described in the previous section (Definition 5). The second set of tests are relative, comparing the stability of different types of networks.

In the descriptive analysis we focus on the simplest statistic capturing the stability of a network, the duration of a network. We measure the duration of a network as the number of payments that occur between when the network is formed and when a different network becomes the basis for payment. We regard the period of time between payments as non-binding negotiation, so, for example, if a link is broken and reformed between payments, we do not consider the second network to be new. We refer to the number of consecutive payments of a given network as its duration in periods.

The duration and frequencies of each type of network in each game are displayed in Figures 5 and 6. To accommodate the large number of networks of short duration while keeping visible differences in the distributions of more stable networks, the data is split between durations of ≤ 5 and > 5 periods.

The complete network in G1 and 1-link in G2 (the PWS networks) clearly display the greatest stability, often lasting upwards of 15 seconds. The 1-link network in G1 (identified by all the farsighted concepts, but not PWS) and the complete network in G2 (identified by only CPS and LPWC, but not PWS) also often last more than half the periods of a game, but much less frequently than the PWS networks. By contrast, distributions of the 2-link networks (identified by PWFS in G1 and LPWC in G2) do not possess these long tails, and there is only one such network that lasts longer than half the game. The empty networks occur rarely, and seldom last more than two seconds.

\[35\] While being a clear indicator of stability, the average duration of networks cannot be used in a formal test of stability due to censoring of final networks in every game. This repeats the remark in footnote 33.
Figure 5: Duration and frequency of networks (Game 1)
Figure 6: Duration and frequency of networks (Game 2)
We turn now to formal statistical tests. All non-parametric tests in this paper are one-tailed binomial tests with eight independent observations. Thus, a hypothesized relationship is significant at the 1% level if it holds true for all eight matching groups in a treatment, and at the 10% level if only for seven. We also test stability using one-sided t-tests, with standard errors clustered by matching group.

We begin by testing whether each type of network is stable according to Definition 5 (Section 3.2). The proportions of paid networks of each type that are again paid in at least the subsequent period, and tests’ results indicating whether these proportions are significantly greater than 0.5, are shown in Table 1. With probabilities of remaining in the same network of 0.92 and 0.89 respectively, the complete network in G1 and 1-link in G2 are found to be stable at the 1% level using both binomial tests and t-tests. The equivalent probabilities for the 1-link network in G1 and the complete network in G2 are 0.68 and 0.79. Neither is found to be stable according to binomial tests, however t-tests find both these figures greater than 0.5, the first at the 5% level, the second at the 1% level.

Both empty networks and the 2-link network in G1 are less likely to remain in the same network than to leave, so cannot be stable according to our definition. The probability of remaining in the 2-link network in G2 is 0.52, but this is not significantly greater than 0.5 according to either the binomial test \( (p = 0.14) \) or t-test \( (p = 0.3) \).

<table>
<thead>
<tr>
<th></th>
<th>Game 1</th>
<th>Game 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Proportions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empty</td>
<td>0.33</td>
<td>0.29</td>
</tr>
<tr>
<td>One-link</td>
<td>0.68</td>
<td>0.89</td>
</tr>
<tr>
<td>Two-link</td>
<td>0.45</td>
<td>0.52</td>
</tr>
<tr>
<td>Complete</td>
<td>0.92</td>
<td>0.79</td>
</tr>
<tr>
<td><strong>Binomial</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empty</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>One-link</td>
<td>0.14</td>
<td>&lt; 0.01***</td>
</tr>
<tr>
<td>Two-link</td>
<td>0.86</td>
<td>0.52</td>
</tr>
<tr>
<td>Complete</td>
<td>&lt; 0.01***</td>
<td>0.79</td>
</tr>
<tr>
<td><strong>t-test</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empty</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>One-link</td>
<td>0.013**</td>
<td>0.14</td>
</tr>
<tr>
<td>Two-link</td>
<td>0.98</td>
<td>0.30</td>
</tr>
<tr>
<td>Complete</td>
<td>&lt; 0.01***</td>
<td>&lt; 0.01***</td>
</tr>
</tbody>
</table>

Table 1: Proportion of paid networks that are also paid in the subsequent period, and p-values of the null hypothesis that this proportion is less than or equal to 0.5.

Comparing the set of networks that are empirically stable according to these tests to the sets identified by different theoretical stability concepts, we can see that only the CPS set identifies them precisely: none of the PWMS, PWFS, vN-MFS, and Level-K stable sets contain the stable complete network in G2, while the LPWC set identifies the unstable 2-link network in G2. Additionally, the
PWMS set does not include the stable 1-link network from G1, and the PWFS contains the unstable 2-link network in G1.

We turn now to comparative tests. The binomial test finds that in both games the PWS networks are more stable than the two-link network and at the 1% level. In G1, the PWS network is also more stable than the 1-link network at the 1% level. In each game, the empty networks are significantly less stable than all other networks. There are no further significant results according to the binomial test.

To take advantage of variation within matching groups, we use a probit regression of the probability of remaining in the current network on the different types of networks, with the 2-link network as the comparison group. Table 2 presents the results of three regressions, the first column using data only from G1, the second from G2, and the third pooling the data from both games. Because in the final regression we want a variable for all pairwise-stable networks, the variable PWS refers to the complete network in G1 and the 1-link networks in G2. We treat the two far-sightedly stable networks separately (FS1 is 1-link network in G1 and FS2 is the complete network in G2).

As can be seen from the first column of Table 2, the networks in G1 can be ranked from least to most stable as empty, 2-link, 1-link, complete, with all relationships significant at the 1% level. The second column reports the results for G2, ranking the networks from least to most stable as empty, 2-link, complete, 1-link, with all relationships significant at the 1% level apart from the last, with the 1-link network being only weakly more stable than the complete ($p = 0.068$).

The last column of Table 2 addresses the possibility that the stability of the complete network is due in part to the fact that it represents an equal share of the surplus to all players, and thus may be focal, or appeal to subjects with fairness concerns. To test this we pool the data from the two games, and include as a regressor a dummy for being a complete network. This controls for fairness concerns because the coefficient on the dummy variable for the complete network in G2 (FS2) now represents extra stability given by factors other than factors shared with the complete network in G1. As can be seen from the final column of Table 2, the coefficient on the complete network dummy is insignificant and close to zero, whereas all the findings from the previous two regressions are unchanged. We are therefore confident that the stability of the complete networks are not due to focality or fairness concerns.\footnote{Unfortunately it is not possible to make one of the asymmetric networks uniquely identified by CPS without either increasing the number of players or giving up anonymity, thus substantially increasing the complexity of the game.}
VARIABLES stay stay stay
Empty -0.0902*** -0.197*** -0.143***
       (0.0246) (0.0519) (0.0314)
PWS 0.446*** 0.351*** 0.371***
       (0.0191) (0.0311) (0.0374)
FS1 0.131*** 0.110***
       (0.0189) (0.0244)
FS2 0.150*** 0.130***
       (0.0362) (0.0421)
Complete 0.0470
         (0.0527)

Observations 13,920 13,920 27,840

PWS: Complete network in G1 and 1-link network in G2;
FS1: 1-link network in G1; FS2: Complete network in G2.
Standard errors clustered by matching group in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 2: Probability of remaining in current network

5 Conclusion

In this paper we propose a new concept of farsighted pairwise stability for a network formation game where players are farsighted, cautious and may care not only about their immediate or long-run payoffs but also about payoffs at intermediate steps. We consider the environment where at least one of full communication or commitment is not possible, and define cautiousness in the spirit of max-min strategies: players will not add or delete a link if there is a possibility that it will make them worse off in the long run. Admittedly, such “extreme pessimism” is appropriate in some but not all network formation games. For example, it is more reasonable in games without too large differences in payoffs. We adopt this approach, as it seems to be the simplest way of capturing cautiousness, without having to deal with beliefs and weighting of a (potentially infinite) number of different alternatives.

We call a set of networks cautious path stable (CPS) if it is the minimal set that satisfies internal and external stability. Namely, a set of networks G is cautious path stable if (1) for any pair of networks in the set, there is no surely improving path (relative to G) between them, (2) from any network outside the set, there exists a surely improving path (relative to G) leading to some network in the set, and (3) no proper subset of G satisfies conditions (1) and (2). The key features underlying this definition – players’ cautiousness and consideration of intermediate payoffs – distinguish the concept of cautious
path stable set from other concepts of farsighted pairwise stability.

We show that a cautious path stable set of networks always exists and provide simple sufficient conditions for a set to be cautious path stable and the unique cautious path stable set. Furthermore, in the setting where players care only about their end-of-path payoffs, we identify some relationships between our concept, which in this setting we refer to as cautious final-network stable set, and the existing farsighted stability concepts such as pairwise farsightedly stable set (PWFS) and von Neumann-Morgenstern pairwise farsightedly stable set (vN-MFS). First, we provide an alternative characterization of a cautious final-network stable set in terms of conditions that appear to be close counterparts of the conditions defining a PWFS set. However, the important difference between the two definitions is that the external stability in our definition requires the existence of not just an improving but surely improving path from any network outside $G$ to a network in $G$, which “adds cautiousness” to players behavior relative to what is assumed by PWFS. Using this result, we then find that any cautious final-network stable set contains at least one PWFS set as a subset, and if a PWFS set is unique, in which case it is also the unique vN-MFS set, then it is a subset of any cautious final-network stable set. A more general reverse statement is not true, as there may exist multiple PWFS and vN-MFS sets that are not included in any cautious final-network stable set.

From our two experimental games we identify four networks as being empirically “stable”, in the sense that once such a network has eventuated, it is more likely than not to remain for the next period: the 1-link and complete networks in each game. Whether or not one agrees with our formal definition of stability, the distributions of durations of these networks are clearly qualitatively different from the others, often lasting at least half the game. Additionally, these networks are significantly more stable than all other networks. CPS is the only theoretical concept which predicts this set exactly, with other concepts either failing to predict an empirically stable network, or predicting a network which is not empirically stable.

The PWS networks are clearly more stable than the networks that are identified only by farsighted stability concepts. Overall our results suggest that farsighted behaviour exists, but is less common than myopic behaviour.
References


A Proofs

Proof of Lemma 1. Suppose \( P = \{g_1, \ldots, g_K\} \) and \( P' = \{g_K, \ldots, g_{K+1}\} \), where \( g_1 = g, \) \( g_K = g' \) and \( g_{K+1} = g'' \). Let \( P \in P^{SI}(g_1, G) \) and \( P' \in P^{SI}(g_K, G') \), where \( G \cap G' \neq \emptyset \) and \( g_{K+1} \in G \). By definition, \( P'' = P \cup P'_2 = \{g_1, \ldots, g_K, g_{K+1}, \ldots, g_{K+N}\} \). Below we will show recursively that for any \( k \) in the decreasing sequence \( K - 1, K - 2, \ldots, 1 \), the continuation of path \( P'' \) from step \( k \), \( P''_k \), is a surely improving path relative to set \( G'' \), where \( G'' \) is any subset of \( G \cap G' \). Then as \( P''_1 = P'' \), the last step of the argument will complete the proof.

Consider \( P''_{K-1} = \{g_{K-1}, g_K, g_{K+1}, \ldots, g_{K+N}\} = \{g_{K-1}\} \cup P' \). Suppose that \( i \) and \( j \) are the players involved in the first-step change on this path, from \( g_{K-1} \) to \( g_K \), i.e., \( g_K = g_{K-1} + ij \) or \( g_K = g_{K-1} - ij \). To show that \( P''_{K-1} \in P^{SI}(g_{K-1}, G'') \), let us first verify that \( P''_{K-1} \in P^{I}(g_{K-1}) \). This follows from the fact that \( P' \in P^{I}(g_K) \) by definition, and players \( i, j \) prefer path \( P' \) to staying in \( g_{K-1} \) for \( |P'| \) steps. The latter is an immediate implication of the fact that \( P \) is a surely improving path relative to \( G \), so that by definition, for any \( \widetilde{P} \in P^{I}(g_K) \) leading to \( G \), including the path \( P' \), the following inequalities hold: (a) \( \pi_i(\widetilde{P}) \geq \pi_i(g_{K-1}^\widetilde{P}) \) and \( \pi_j(\widetilde{P}) \geq \pi_j(g_{K-1}^\widetilde{P}) \), with at least one inequality being strict, if \( g_K = g_{K-1} + ij \), or (b) \( \pi_i(\widetilde{P}) > \pi_i(g_{K-1}^\widetilde{P}) \) if \( g_K = g_{K-1} - ij \). Now, given that \( P' \) is a surely improving path relative to \( G' \) and hence, also relative to \( G'' \subseteq G' \), that is, \( P' \in P^{SI}(g_K, G'') \), and inequalities (a), (b) hold for any \( \widetilde{P} \in P^{I}(g_K) \) that leads to \( G \) and hence, also for any improving path that leads to \( G'' \subseteq G \), it follows that conditions (i) and (ii) of the definition of a surely improving path relative to \( G'' \) are satisfied for all steps on the path \( P''_{K-1} = \{g_{K-1}\} \cup P' \). Thus, \( P''_{K-1} \in P^{SI}(g_{K-1}, G'') \).

Next, consider \( P''_{K-2} = \{g_{K-2}, g_K, g_{K-1}, g_K, \ldots, g_{K+N}\} = \{g_{K-2}\} \cup P''_{K-1} \). Repeating the same argument as before, we will conclude that \( P''_{K-2} \in P^{SI}(g_{K-2}, G'') \). Then by analogy, we can construct a sequence of surely improving paths \( P''_{K-1}, P''_{K-2}, P''_{K-3}, \ldots, P''_2, P'' \). Thus, \( P'' \in P^{SI}(g_1), \) where \( g_1 = g \).

Proof of Proposition 2. Below we propose an algorithm to construct one cautious path stable set.

1. Consider the whole network space \( G \). \( G = G \) trivially satisfies condition (2) of the definition of a cautious path stable set. If it also satisfies condition (1), then either \( G \) is a cautious path stable set – when \( G \) is the minimal set that satisfies (1) and (2), or there exists \( G' \subset G \) that is cautious path stable. In the latter case, the existence of a cautious path stable set is guaranteed, as if \( G \) is not minimal, there must exist \( G' \subset G \) that satisfies conditions (1) and (2). Then either \( G' \) is the
minimal set that satisfies (1) and (2), so that $G'$ is cautious path stable, or there exists a proper subset of $G'$ that satisfies both conditions, etc. As the cardinality of set $G'$ is finite, the sequence of thus constructed subsets of $G'$ that satisfy (1) and (2) is finite, and the last, “smallest” subset in this sequence is minimal, that is, satisfies all three conditions of the cautious path stable set.

Now, suppose that $\mathcal{G}$ does not satisfy condition (1). Then there must exist a network $g \in \mathcal{G}$ and path $\tilde{P} \in P^{SI}(g, \mathcal{G})$ such that $\tilde{P}$ leads to $\mathcal{G} \setminus \{g\}$.

2. Consider $G_1 = \mathcal{G} \setminus \{g\}$. $G_1$ trivially satisfies condition (2). If it also satisfies condition (1), then either $G_1$ or one of its proper subsets is a cautious path stable set. If condition (1) is not satisfied, then there exist a network $g_1 \in G_1$ and path $\tilde{P}' \in P^{SI}(g_1, G_1)$ such that $\tilde{P}'$ leads to $G_1 \setminus \{g_1\}$.

3. Consider $G_2 = G_1 \setminus \{g_1\}$. $G_2$ satisfies condition (2). To verify that, suppose that condition (2) is not satisfied. Since from $g_1$ there exists a path $\tilde{P}'$ leading to $G_2$ that is surely improving relative to $G_2$, it must be that there does not exist a surely improving path relative to $G_2$ from $g$ to $G_2$. But from $g$ there exists a path $\tilde{P}$ leading to $G_1 = \mathcal{G} \setminus \{g\}$ that is surely improving relative to $G_1$ and hence, also surely improving relative to $G_2$. Then it must be that any such surely improving path $\tilde{P}$ leads to network $g_1$. So, there exist two surely improving paths: $\tilde{P} \in P^{SI}(g, G_2)$ that leads to $g_1$ and $\tilde{P}' \in P^{SI}(g_1, G_2)$ that leads to $G_2$. Then by Lemma 1, path $\tilde{P} \cup \tilde{P}' \in P^{SI}(g, G_2)$ and it leads to $G_2$. This contradicts the assumption that there does not exist a surely improving path relative to $G_2$ from $g$ to $G_2$. Thus, $G_2$ satisfies condition (2).

If $G_2$ also satisfies condition (1), then either $G_2$ or its proper subset is a cautious path stable set. If condition (1) is not satisfied, then we reduce the set even further by constructing $G_3 = G_2 \setminus \{g_2\}$, etc.

Iterating this reasoning, we can build a decreasing sequence $\{G_k\}_{k \geq 1}$ of proper subsets of $\mathcal{G}$, satisfying condition (2). As $\mathcal{G}$ has a finite cardinality, and as a set consisting of a single network trivially satisfies condition (1), there exists $K \geq 1$ such that $G_K \neq \emptyset$ and satisfies both conditions, (1) and (2). Then following the same logic as before, either $G_K$ also satisfies condition (3) of minimality and hence, it is a cautious path stable set, or there exists a proper subset of $G_K$ that satisfies all three conditions. This completes the proof. ■

Proof of Proposition 5.

37 Recall that one property of surely improving paths implies that $P^{SI}(g, \mathcal{G}) \subseteq P^{SI}(g, G_1)$. 
Let $G$ be CFNS set. Let us verify that conditions (i), (ii) and (iii) of Proposition 5 hold. In fact, it is enough to verify that conditions (i) and (ii) hold, as then (iii) is satisfied, too. Indeed, if (iii) is not satisfied, then there exists a proper subset of $G$, $G' \subset G$, such that (i) and (ii) hold for $G'$. Consider the minimal among such subsets, i.e., $G' \subset G$ that satisfies all three conditions, (i), (ii) and (iii). But then from the proof of sufficiency ($\Leftarrow$) it follows that $G'$ must satisfy conditions (1) and (2) of a CFNS set, which contradicts the minimality of the CFNS set $G$.

So, let us focus on conditions (i) and (ii). Clearly, condition (ii) follows immediately from the definition of a CFNS set. Suppose condition (i) does not hold. This means that at least one of the two statements, (a) or (b), is true:

(a) $\exists g \in G$ and $ij \notin g$ such that $g + ij \notin G$, and $\forall g' \in F^I(g + ij) \cap G$ it holds that $(Y_i(g'), Y_j(g')) > (Y_i(g), Y_j(g));$  

(b) $\exists g \in G$ and $ij \in g$ such that $g - ij \notin G$, and $\forall g' \in F^I(g - ij) \cap G$ it holds that $Y_i(g') > Y_i(g).$  

If (a) is true, then the inequality $(Y_i(g'), Y_j(g')) > (Y_i(g), Y_j(g))$ holds, in particular, for $\tilde{g} \in F^{SI}(g + ij, G) \cap G$. Such network $\tilde{g}$ exists, as $F^{SI}(g + ij, G) \cap G \neq \emptyset$ due to condition (2) of the definition of a CFNS set. This, together with the fact that $(Y_i(g'), Y_j(g')) > (Y_i(g), Y_j(g))$ for any $g' \in F^I(g + ij) \cap G$, means that $F^{SI}(g, G) \cap G \neq \emptyset$. However, this contradicts condition (1) of the definition of a CFNS set.

Similarly, if (b) is true, then the inequality $Y_i(g') > Y_i(g)$ holds, in particular, for $\tilde{g} \in F^{SI}(g - ij, G) \cap G$. As before, such network $\tilde{g}$ exists due to condition (2) of the definition of a CFNS set. This, together with the fact that $Y_i(g') > Y_i(g)$ for any $g' \in F^I(g - ij) \cap G$, means that $F^{SI}(g, G) \cap G \neq \emptyset$. However, this contradicts condition (1) of the definition of a CFNS set.

Thus, neither (a) or (b) holds, hence, condition (i) is satisfied.

($\Leftarrow$): Suppose that set $G$ is such that conditions (i), (ii) and (iii) of Proposition 5 hold. Let us verify that $G$ is a CFNS set, that is, satisfies conditions (1), (2) and (3). In fact, it is enough to verify conditions (1) and (2), as then (3) follows. Indeed, if not, then there must exist a proper subset of $G$, $G' \subseteq G$, such that $G'$ satisfies (1) and (2). But from the proof of necessity ($\Rightarrow$) we know that conditions (1) and (2)

---

38Such minimal subset of $G$ exists as otherwise we could construct an infinite declining sequence of subsets of $G$, all satisfying conditions (i) and (ii). This, however, contradicts the fact that $G$ has a finite cardinality.

39We use the notation $(Y_i(g'), Y_j(g')) > (Y_i(g), Y_j(g))$ for $Y_i(g') \geq Y_i(g)$ and $Y_j(g') \geq Y_j(g)$ with at least one inequality holding strictly.

40Note that this inequality holds for one and the same player, $i$ or $j$. That is, given link $ij$, there exists one player, $i$ or $j$, such that her payoff in $g'$ is larger than in $g$ for $\forall g' \in F^I(g - ij) \cap G$. Otherwise, (b) would not be a contradiction to condition (ib).
imply (i) and (ii), that is, a proper subset of $G$, $G'$, must satisfy (i) and (ii). This, however, contradicts the minimality of set $G$ established by condition (iii).

Let us focus on conditions (1) and (2). Condition (2) is trivially satisfied, as it is identical to (ii). If condition (1) is also satisfied, then the proof is completed. Note that this is trivially the case when $G$ consists of a single network. Suppose now that set $G$ contains at least two networks, i.e., $|G| \geq 2$, and condition (1) is not satisfied. This means that $\exists g \in G$ such that $F^{SI}(g,G) \cap G \neq \emptyset$. We claim that this violates condition (iii) of minimality in Proposition 5.

Claim: There exists $G' \subseteq G$ that satisfies conditions (i) and (ii).

Below we construct this set $G'$. Consider $G_1 = G \setminus \{g\}$. Note that $|G_1| \geq 1$ as $|G| \geq 2$. $G_1$ satisfies condition (ii). Indeed, suppose that it doesn’t. Since $F^{SI}(g,G_1) \cap G_1 \supseteq F^{SI}(g,G) \cap G \neq \emptyset$, it must be that for some $g' \in G \setminus G$, $F^{SI}(g',G_1) \cap G_1 = \emptyset$. On the other hand, as $G$ satisfies condition (ii), $F^{SI}(g',G) \cap G \neq \emptyset$, and since $F^{SI}(g',G_1) \supseteq F^{SI}(g',G)$, we have $F^{SI}(g',G_1) \cap G \neq \emptyset$. Together, $F^{SI}(g',G_1) \cap G \neq \emptyset$ and $F^{SI}(g',G_1) \cap G_1 = \emptyset$, mean that $F^{SI}(g',G_1) \cap G = \{g\}$. So, we have $F^{SI}(g',G_1) \cap G = \{g\}$ and $F^{SI}(g',G_1) \cap G_1 \neq \emptyset$, which by Lemma 1 implies the existence of a surely improving path relative to $G_1$ from $g'$ to $G_1$, i.e., $F^{SI}(g',G_1) \cap G \neq \emptyset$. But this contradicts the assumption about $g'$. Hence, $G_1$ satisfies condition (ii).

Now, if $G_1$ also satisfies condition (i), then the proof is completed. Note that this is trivially the case when $G_1 = \{g_1\}$, that is, consists of a single network. Indeed, in this case, (i) is satisfied as for any $i, j$, $g_1 \pm i j \in G \setminus G_1$ and by condition (ii), there exists a surely improving path relative to $G_1$ from $g_1 \pm i j$ that leads back to $g_1$, i.e., $F^{SI}(g_1 \pm i j, G_1) \cap G_1 = \{g_1\}$. As payoffs of $i$ and $j$ in the end of this path are equal to their payoffs in $g_1$, all pairwise deviations from $g_1$ are deterred.

So, suppose that $G_1$ contains at least two networks, i.e., $|G_1| \geq 2$, and condition (i) is not satisfied. This means that at least one of the two statements, (a) or (b), is true:

(a) $\exists g_1 \in G_1$ and $i j \notin g_1$ such that $g_1 + i j \notin G_1$, and $\forall g'_1 \in F^I(g_1 + i j) \cap G_1$ it holds that $(Y_i(g'_1), Y_j(g'_1)) > (Y_i(g_1), Y_j(g_1))$;

(b) $\exists g_1 \in G_1$ and $i j \in g_1$ such that $g_1 - i j \notin G_1$, and $\forall g'_1 \in F^I(g_1 - i j) \cap G_1$ it holds that $Y_i(g'_1) > Y_i(g_1)$.

In particular, the above is true for $\tilde{g} \in F^{SI}(g_1 \pm i j, G_1) \cap G_1$, which exists due to the fact that $G_1$ satisfies (ii). This, together with the fact that the payoffs of $i$ and $j$ improve at any $g'_1 \in F^I(g_1 \pm i j) \cap G_1$ (i.e., the inequalities hold for any $g'_1 \in F^I(g_1 \pm i j) \cap G_1$), means that $F^{SI}(g_1, G_1) \cap G_1 \neq \emptyset$.\"
Let us define $G_2 = G_1 \setminus \{g_1\}$. $|G_2| \geq 1$ as $|G_1| \geq 2$. Repeating the same argument as before, but with respect to $G_2$ instead of $G_1$, we can show that $G_2$ satisfies condition (ii). If it also satisfies condition (i), then the proof is completed; otherwise, we construct $G_3$, etc. Iterating this reasoning, we can construct a decreasing sequence \{\text{G}_k\}_{k \geq 1} of proper subsets of $G$, each satisfying condition (ii). As $G$ has a finite cardinality, and as a set consisting of a single network trivially satisfies condition (i), there exists $K \geq 1$ such that $G_K \neq \emptyset$ and satisfies both conditions, (i) and (ii). Denoting this set $G_K$ by $G'$, we complete the proof of the claim and the proof of the proposition.

**Proof of Proposition 9.** The proof requires defining the sequence of sets \{\text{Z}_k\}_{k \geq 1} used to find the LPWC set. The iterative procedure to construct this sequence and find the largest consistent set was proposed by Chwe (1994) and is described in Herings et al. (2009, p. 539). Let $Z^0 \equiv \emptyset$. Then $Z^k (k = 1, 2, \ldots)$ is inductively defined as follows: $g \in Z^{k-1}$ belongs to $Z^k$ with respect to $Y$ if

(ia) \forall ij \notin g \exists g' \in Z^{k-1}$, where $g' = g + ij$ or $g' \in F^I(g + ij)$ such that $(Y_i(g'),Y_j(g')) = (Y_i(g),Y_j(g))$ or $Y_i(g') < Y_i(g)$ or $Y_j(g') < Y_j(g)$,

(ib) \forall ij \in g \exists g',g'' \in Z^{k-1}$, where $g' = g - ij$ or $g' \in F^I(g - ij)$, and $g'' = g - ij$ or $g'' \in F^I(g - ij)$, such that $Y_i(g') \leq Y_i(g)$ and $Y_j(g'') \leq Y_j(g)$.

The LPWC set is given by $\bigcap_{k \geq 1} Z^k$.

Note that as \{\text{g}\} is a CFNS set, it holds by definition that for all $ij \notin g$, $g \in F^S^I(g + ij,\{g\})$, so that, in particular, $g \in F^I(g + ij)$. Similarly, for all $ij \in g$, $g \in F^S^I(g - ij,\{g\})$, so that $g \in F^I(g - ij)$.

This means that $g \in Z^1$. By induction, $g \in Z^k$ for all $k \geq 1$. Thus, $g$ belongs to the LPWC set.

**Proof that in Game 1 there are no improving paths from g7 to any other network.** Consider all possibilities in turn. Notice that any improving path to the empty network must pass via a 1-link network at the previous step. But the last step from a 1-link network to the empty network is not increasing the payoff of a player who deletes the link, hence, it cannot be the last step of any improving path. Similarly, in order to reach a 1-link network one must pass via a 2-link network or the empty network at the previous step, where the empty network must itself be preceded by some 1-link network. But the last step from the 2-link network to the 1-link network is not an improving path (32 < 30 and 17 < 6), and neither is the two-step path 1-link $\rightarrow$ empty $\rightarrow$ 1-link ($\frac{1}{2}(5 + 30) < 30$). Hence, those last steps cannot belong to any improving path. Finally, reaching a 2-link network requires passing via
a 1-link or the complete network at the previous step, where a 1-link network must itself be preceded by either a 2-link network or the empty network. However, the last step from the complete to a 2-link network is not improving \((17 < 22)\) and neither is the two-step path \(2\text{-link} \rightarrow 1\text{-link} \rightarrow 2\text{-link}\) \(\frac{1}{2}(30 + 32) < 32, \frac{1}{2}(30 + 17) < 32\) and \(\frac{1}{2}(6 + 17) < 17\). Similarly, while a path \(\text{empty} \rightarrow 1\text{-link} \rightarrow 2\text{-link}\) is improving, a longer path including the preceding step, \(1\text{-link} \rightarrow \text{empty} \rightarrow 1\text{-link} \rightarrow 2\text{-link}\) is not improving, as in the best case, a player in the 1-link network who initiates a move on this path receives \(\frac{1}{3}(5 + 30 + 32) < 30\). Thus, we ruled out all possibilities of an improving path from the complete network. ■

B Experimental results

<table>
<thead>
<tr>
<th></th>
<th>Empty</th>
<th>One-link</th>
<th>Two-link</th>
<th>Complete</th>
</tr>
</thead>
<tbody>
<tr>
<td>Game 1</td>
<td>0.33</td>
<td>0.68</td>
<td>0.45</td>
<td>0.92</td>
</tr>
<tr>
<td>Matching group 1</td>
<td>0.33</td>
<td>0.91</td>
<td>0.51</td>
<td>0.99</td>
</tr>
<tr>
<td>Matching group 2</td>
<td>0.27</td>
<td>0.66</td>
<td>0.44</td>
<td>0.83</td>
</tr>
<tr>
<td>Matching group 3</td>
<td>0.37</td>
<td>0.71</td>
<td>0.51</td>
<td>0.89</td>
</tr>
<tr>
<td>Matching group 4</td>
<td>0.28</td>
<td>0.91</td>
<td>0.40</td>
<td>0.98</td>
</tr>
<tr>
<td>Matching group 5</td>
<td>0.43</td>
<td>0.40</td>
<td>0.47</td>
<td>0.88</td>
</tr>
<tr>
<td>Matching group 6</td>
<td>0.37</td>
<td>0.70</td>
<td>0.46</td>
<td>0.98</td>
</tr>
<tr>
<td>Matching group 7</td>
<td>0.10</td>
<td>0.44</td>
<td>0.47</td>
<td>0.93</td>
</tr>
<tr>
<td>Matching group 8</td>
<td>0.33</td>
<td>0.55</td>
<td>0.38</td>
<td>0.65</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Empty</th>
<th>One-link</th>
<th>Two-link</th>
<th>Complete</th>
</tr>
</thead>
<tbody>
<tr>
<td>Game 2</td>
<td>0.29</td>
<td>0.89</td>
<td>0.52</td>
<td>0.79</td>
</tr>
<tr>
<td>Matching group 9</td>
<td>0.40</td>
<td>0.92</td>
<td>0.51</td>
<td>0.42</td>
</tr>
<tr>
<td>Matching group 10</td>
<td>0.22</td>
<td>0.88</td>
<td>0.62</td>
<td>0.39</td>
</tr>
<tr>
<td>Matching group 11</td>
<td>0.31</td>
<td>0.91</td>
<td>0.58</td>
<td>0.51</td>
</tr>
<tr>
<td>Matching group 12</td>
<td>0.30</td>
<td>0.98</td>
<td>0.53</td>
<td>0.54</td>
</tr>
<tr>
<td>Matching group 13</td>
<td>0.37</td>
<td>0.77</td>
<td>0.49</td>
<td>0.86</td>
</tr>
<tr>
<td>Matching group 14</td>
<td>0.15</td>
<td>0.89</td>
<td>0.54</td>
<td>0.92</td>
</tr>
<tr>
<td>Matching group 15</td>
<td>0.22</td>
<td>0.89</td>
<td>0.59</td>
<td>0.79</td>
</tr>
<tr>
<td>Matching group 16</td>
<td>0.27</td>
<td>0.75</td>
<td>0.34</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Table 3: Probability of remaining in current network
C Instructions

[ONSCREEN]

Before the experiment begins there will be a short tutorial and three practice games to make sure everybody understands how points can be earned.

The points that are distributed in these three practice rounds will not affect your final payment.

Please click "Continue" to proceed to the tutorial.

[PRINTED]

Please read and follow these instructions. Text in italics describes things you should do onscreen. If you have a question, raise your hand and someone will come to help you as soon as possible.

- In each round of this experiment you will be interacting with two other people using the screen you can see on your monitor.
- You are represented by the green dot, and the other two players by the blue dots.
- Links may be formed between two players in the following way:
  - You can indicate that you are willing to form a link with another person by clicking on their blue dot.
  - Clicking on them again indicates you are no longer willing to form a link with them.
  - You can click on a person as many times as you like, switching back and forth between being willing to form a link with them or not.
  - If two people have both clicked on each other then a link is formed and it is shown in red.
  - A link can be formed between to people only if both of them want it to be formed.
  - If only one of the people has shown they are willing to form a link then it is shown in pink.
- On the screen in front of you the two other people have formed a link, and the person on your right has indicated they are willing to form a link with you.
  - Click on the blue dot on the left and see how the line turns pink. Click again and see how it becomes white again.
  - Click on the blue dot on the right and see how the line turns red. Click again and see how it becomes pink again.
  - Notice that nothing you can do will change the colour of the line between the other two people. Whether or not that link is formed depends only on their decisions.
• Every second, you and the two other people will earn points. The number of points per second earned by each person is shown in red next to their dot.
• The number of points each person earns per second depends on which links are formed at that point in time.
• These numbers will vary from round to round and will be shown to you before each round begins. The numbers for the screen you see in front of you are described in the following diagram:
• Click on the two other people and see how the numbers in red change, and how they relate to the diagram. Notice that it doesn’t make a difference if a line is white or pink; the numbers change only if a link is formed or broken (i.e. becomes red, or changes from red to pink).
• In the practice rounds and real interactions the screen will look slightly different. An example is shown below:

![Diagram](image)

• As mentioned before, points will be earned every second. The total number of points you have earned so far in a round will be shown at the top left of the screen as shown in the picture “Your Point”). This number will increase every second by the red number below the green dot.
• Each round lasts for 30 seconds. The number of seconds left will be shown at the top right.
• WHEN YOU HAVE UNDERSTOOD THESE INSTRUCTIONS, PLEASE CLICK THE BUTTON ON YOUR COMPUTER SCREEN.

[ONSCREEN]

Please answer the following questions relating to the picture shown in the handout.

Click "Continue" when you have answered all questions.

How many points are you earning per second?
How many points is the person on your left earning per second?
How many points have you earned so far this round?
How many seconds are left before the round ends?

[New Screen]

You have answered all the questions correctly.

Before the real experiment begins there will be three practice rounds.
These practice rounds will not affect your final payment.

The purpose of these practice rounds is for you to learn how these interactions work. You should use them to experiment and learn how links are formed and how they relate to the payoffs. Do not worry about earning a lot of points because they do not count!

The points associated with each practice round are shown on your handout.

If you have any questions, please raise your hand and someone will come to help you as soon as possible.

Otherwise please click "Continue" and wait for the other participants to finish the Tutorial.

The first practice round is about to begin.

Check the payoffs described in Figure 1 of your printed instructions. These are the payoffs that are relevant for the practice rounds.

Click OK when you are ready to begin.

[New Screen]

The first practice round is about to begin.

Check the payoffs described in Figure 1 of your printed instructions. These are the payoffs that are relevant for the practice rounds.

[New Screen]

Click OK when you are ready to begin.
The practice rounds are now over.

You will now be handed the diagrams which describe the payoffs for the first real rounds.

[New Screen]

You will now play a game similar to the one in the tutorial but with different payoffs.

Please look at the diagram you have just been given to see how the points you earn will depend on the links that are formed.

You will play this game 20 times. The links that have been already formed at the beginning of the game will be randomly determined each time.

After each time you will be randomly rematched with new participants. This means it is unlikely you will be playing with exactly the same people as in the previous round.

When all games have been played, one game will be randomly chosen to determine how much you will be paid. All participants will be paid for the same game. For every 90 points you earn in that game you will be paid 2 Euros.

When you are ready to start, please click "Continue".