On the sequential formation of networks

Liza Charroin

On the sequential formation of networks*

Liza Charroin

June 23, 2014

Abstract

The network formation model described in this research paper is based on the connection model of Bala and Goyal [1]. We extend this model by introducing two properties: heterogeneity of players and sequentiality of the network formation process. These two characteristics modify the emergence of equilibrium and efficient networks. Our main theoretical finding is that sequentiality is the main determinant of the equilibrium network while heterogeneity is secondary. The equilibrium often requires that the last player of the game incurs a large part of the linking costs. On the contrary, the leader does not create links at the equilibrium. In the following of the paper we build an experimental protocol. Our purpose is to analyse whether our model predicts behaviour in the laboratory and whether individuals are more attracted by the equilibrium or by efficiency.

1 Introduction

Networks are prominent in many real life examples. The success of Facebook, Linked In or other Web social networks shows the increasing importance of networks in our society. They have also been studied by many researchers to understand market sharing agreements [2], job search [6], friendship [11] or co-authorship [22] and [24]. This is a very wide field that allows to comprehend numerous situations. In our paper we study social and economic networks theoretically and experimentally. Jackson [21] wrote a theoretical survey about network formation that shows the great variety of models. Kosfeld [23] also made a survey but from an experimental point of view to

*I thank Marie Claire Villeval and Christophe Bravard for their help for the creation of this research paper. I also wish to thank Yann Bramoullé, Sudipta Sarangi and Jurjen Kamphorst for their precious comments.

show the different applications in the laboratory with human subjects.

A social or economic network is composed by agents who create links. In this literature, links are valuable but also costly. They need to make a tradeoff between the value of a link and its cost. Studying this field allows a better understanding of linking decisions. The purpose is to answer the following questions: who makes the first move to connect with another player? why do people want to link with some agents and not others? why does an agent accept to pay a cost for a connection? To sum up, the central question of this paper is which type of networks emerges and for which reasons.

We base our theoretical model on a seminal paper in the network formation field: the connection model of Bala and Goyal [1] (henceforth BG). In their model players can create links *unilaterally* (without the agreement of the potential partners). The formation of a link is costly and only the initiator of the link pays this cost. Players can form *direct links* but they can also be connected indirectly with other players and benefit from them, i.e. "friends of my friends are also my friends". These players are *indirect links*.

The authors distinguish two models: the one-way flow model and the two-way flow model. In the former, information is only received by the initiator of the link and in the latter, both players benefit from the connection (even if one player did not pay for it). To summarize, in the first model, players benefit only from links they created. In the second model, players benefit from all the connections even if they did not pay for them. In addition, in both models, players benefit from indirect links.

Bala and Goyal want to characterize the *strict Nash* networks. A network is strict Nash if no player has a strict incentive to change strategy if the other players keep the same strategy. The result of the one-way flow model is that each player should create only one link with another player. The architecture of this network looks like a circle. This equilibrium network is called the *wheel*. For the two-way flow model the strict Nash network is the *center-sponsored star* network (CSS) with one *central player* and *peripheral players*. A center-sponsored star is a network where the player in the center forms links with every other player and others form no link. The central player sponsors the whole network. In this network, peripheral players benefit from everybody without paying anything. The main issue with a star network is that one player has to take the particular position of central player. This player has to incur all the linking costs. This position is very disadvantageous and leads to the instability of the network. The central player does not want to stay in this position and has an incentive to remove links.

Our model is based on the second model of BG. We chose the two-way flow model because of the instability and the asymmetry of the equilibrium network that can be studied theoretically and experimentally. Indeed in the laboratory the two-way flow model leads to interesting psychological and sociological conclusions. We complete the research of BG by introducing two characteristics.

The main contribution of our research paper is that we study network formation in a sequential manner. It means that players make their linking decisions one after another. The major part of the literature studies simultaneous network formation, where players can take decisions in the same time. Our paper is inspired by the Stackelberg competition [29] in the sense that one player is the *market leader* and the others are *followers*. The market leader plays first and the others play after him and in a sequential manner. This assumption has been widely used in industrial economics with the example of competing firms. Nevertheless sequentiality can be transposed in network formation models for a more general purpose. Sequentiality brings new perspectives and represents situations where players can analyse the decisions of preceding players. Moreover players that play at the beginning of the formation of the network have more power to orientate the network. It can be applied to many economic situations. Agents may need time to create links and simultaneous decisions are not always possible. For example, before the collaboration of two firms, they analyse the preceding collaborations. It can also explain the formation of co-authorship between scientists. An author can look at the preceding co-authorship of the scientist before creating a link with him.

In addition we study network formation among heterogeneous players. The major part of the literature assumes homogeneity of players. In our setting, players have different characteristics. Galeotti, Goyal and Kamphorst [17] introduced heterogeneity in the two-way flow model of BG. More precisely they fixed arbitrary heterogeneous costs of linking and heterogeneous values of accessing others. The latter means that some players can benefit more from their connections than other players. They have a better capacity to absorb the value of others. Their results are similar than the results of BG in many cases but there is one major difference. Their main finding is that an equilibrium can be asymmetric and stable. A network is *asymmetric* if payoffs and strategies are heterogenous among players. The star is considered as asymmetric because players have different strategies and payoffs are unequal between the central player and peripheral players. In BG the central player has an incentive to remove links. But in the model of Galeotti et al. with the introduction of heterogeneity some natural candidates emerge to take the particular place of the central player (players with a high-value or a low-cost of linking). Consequently the star network is a strict Nash network because no player wants to deviate from his strategy.

Like in their paper, we use heterogeneous costs of linking. However instead of introducing heterogeneous values of accessing others, we study heterogeneous personal values (own value of the player). The profit of a player depends on the value of his partners and the costs he paid to form links. Sequentiality and heterogeneity can be the solution for the instability issue. In our setting players with a low cost of linking or a higher value than the others are natural candidates as central player. To sum up, our framework is based on the two-way flow model and we add two properties: the heterogeneity property of Galeotti et al. and the principle of sequentiality.

The first part of the paper describes our theoretical model with the two properties (sequentiality and heterogeneity). The purpose of our paper is to see the contribution of these particular characteristics in the model. We want to see the changes induced by these two characteristics in the emergence of networks. Our hypothesis is that they will partly solve the instability problem that BG witnessed with their equilibrium network (the star). The main theoretical result of our paper is that sequentiality is more important than heterogeneity in the formation of networks. Indeed sequentiality is the major determinant of the equilibrium. The heterogeneity property is secondary. Globally, the player in last position is often the central player of the CSS at the equilibrium. As a consequence he pays all the linking costs. The preceding players know that this last player will have no choice but creating links if he wants a positive profit. They can remain passive and let the last player sacrifice himself and bear all the linking costs.

In the second part of the paper we are going to build an experimental protocol. We want to test our model in the laboratory to see the differences with the theoretical results. The laboratory allows us to create a "microeconomic environment [...] where adequate control can be maintained and accurate measurement of relevant variables guaranteed" [32]. More precisely, in a laboratory the experimenter can control the characteristics of the players and the different variables of the procedure. Moreover he knows all the potential links that can be formed and the information accessed by each player. Two papers tested the connection model in the laboratory ([3], [13]). The authors witnessed the instability of the equilibrium with the two-way flow model. Their results were not in accordance with the theoretical predictions. To solve this issue, Goeree, Riedl and Ule [19] introduced heterogeneity of players. Thanks to heterogeneity, “natural candidates” emerge to take the particular position of central player (low-cost or high-value player).

In addition in our procedure we introduce *objective* (or *endogenous*) characteristics. In the major part of the literature, the characteristics of the players are exogenous and arbitrary. The experimenter randomly gives characteristics. In our setting, each player reveals his personal value and cost thanks to tests that will be displayed before the network formation game. The fact that players are at the origin of their characteristics should lead to more strategic and self-interested behaviours because they feel legitimate.

Experimentally, our purpose is threefold. First, we want to test the results of the theoretical model. The goal is to analyse if equilibrium networks emerge with our framework. An issue may be coordination problems. We investigate if decentralized decisions (players cannot communicate) can lead to the equilibrium and if agents are able to coordinate without the help of an external authority. However, players do not only follow their self-interest. They may want to improve the well-being of the others and have social motives. More precisely, they can have efficiency and fairness concerns. These concepts describe a preference for a high aggregate payoff and for egalitarian payoffs. Consequently our second purpose is to see the impact of these motives on the emergence of networks. Finally, our last purpose is to test the contribution of objective characteristics instead of exogenous and arbitrary characteristics. We analyse this feature thanks to a control treatment where characteristics will be completely random. The literature shows that if players earn their position of leader, they play more strategically than if it is due to a random assignment. Consequently endogenous characteristics may counter the effect of social motives. The goal is to see how and why endogenous characteristics change the formation of network and the type of network structure that emerges.

In the following of the paper section 2 presents the experimental literature on network formation. In section 3 we describe the theoretical model and its results. Then in section 4 we explain the experimental procedure and we provide our predictions of experimental results in section 5. Finally section 6 concludes.

2 Experimental literature

Only few experiments have been made on network formation. A major part of the experiments deals with network formation where mutual consent is necessary to form a link ([5], [10], [27], [28], [30], [31]). In these experiments both players must agree to form a link and they share the linking cost. Only few experimental papers tested the connection model of Bala and Goyal where networks are directed ([3],[13] and [19]). In these experiments unilateral consent is sufficient to form a link. We focus on these last papers in this literature review as they deal with the model of Bala and Goyal. These three papers are presented in this section and will guide us for our experimental protocol.

These papers are complementary. The first paper is the experiment of Falk and Kosfeld [13]. They tested the original model (both one-way and two-way flow model) with homogeneous values and costs and where players can form links simultaneously. Their main finding is that the connection model predicts outcomes very well with the one-way flow model but it largely fails to predict the data with the two-way flow model. Concretely strict Nash networks emerged in the one-way but not in the two-way flow treatment. They found three reasons to explain this phenomenon: asymmetry, efficiency and fairness. First, it may be due to *asymmetry*. As we said a network is asymmetric if payoffs and strategies are very different across players. The strict Nash equilibrium of the one-way flow model is the wheel and it is the CSS for the two-way flow model. There is no asymmetry with the wheel because each player forms one link and pays for one linking cost. However the star requires a central player that builds the whole network. There may be a coordination problem to find this particular player. Strategies and payoffs are very heterogeneous with the star. However asymmetry does not explain everything. Indeed in some cases (where the cost is high) the strict Nash network is the empty network (no links are created) which is symmetric. Players never played this equilibrium, probably for one reason: *efficiency*. The concern for efficiency describes the fact that players want to increase the aggregate payoff of the network. A network is efficient if it maximizes the total profit of the network. In the experiment of Falk and Kosfeld, players preferred to create links instead of having the empty network even if it was not the equilibrium. Finally the authors developed the idea of *inequality aversion* (fairness). This social preference is very important for the participants. It means that players do not like large gap of payoffs. They prefer homogeneous profits and reject inequalities. With the CSS in the two-way flow model, payoffs are not equal between the central and peripheral players because the central player has to bear all the linking costs. On the contrary for the wheel, payoffs are equal.

Further to the paper of Falk and Koslfeld, Berninghaus, Ott and Vogt [3] tried to understand the mitigated results with the two-way flow model. Like Falk and Kosfeld they used the connection model but focused on the two-way flow model. They introduced some theoretical changes to create a favourable environment for the emergence of equilibrium networks. For that the authors distinguished active, passive and indirect neighbours. If two players i and j are linked and that i is the initiator of the link, j is an *active neighbour* of i. On the contrary, i is a *passive neighbour* of j. Finally, the active or passive neighbours of an active neighbour are *indirect neighbours.* With our example, if i paid to form a link with another player j, he can benefit from the neighbours of j but j cannot benefit from the neighbours of i. This environment changes the theoretical results. Indeed the strict Nash equilibrium is no more the CSS or the empty network, it is the empty network or the *periphery-sponsored star* (PSS) where each peripheral player pays for one link with the central player and the central player remains inactive. This network is sponsored by peripheral players. The experimental results of Berninghaus et al. [3] differ from the results of Falk and Kosfeld. They found that half of the groups reached the strict Nash network (the PSS) or converged to this network. They witnessed a strong learning effect. There are few reasons that can explain the differences between the two types of results. The first reason is complexity. Indeed the star is asymmetric and coordinating on the CSS is even more complicated than coordinating on the PSS. With the CSS, one player needs to link to all other players and peripheral players need to stay inactive. For the PSS only one player needs to understand that he has to remain passive. The other players only have to create one link. The second reason is inequality aversion. Indeed the strategy and payoff inequalities are very large with the CSS while payoffs are quite egalitarian with the PSS.

Even if Berninghaus et al. have interesting results, they do not solve the problem of the CSS. One paper tried to find an environment where the CSS can emerge in the laboratory. It is the paper of Goeree, Riedl and Ule

[19]. The authors noticed that stars rarely emerge in network experiments due to asymmetry, inequalities and instability. However many structures are star-like networks in real life examples with central players and peripheral players who are less active. The purpose of their paper is to show under which conditions a CSS can emerge. They based their model on the two-way flow model of BG [1] with decay (the value of a link is weighted with the distance between the two players). A player benefits from the value of his partners (weighted with distance) but pays a linking cost if he is the initiator of the link. The differences with the two preceding papers is that they introduced heterogeneity of players. They studied treatments where some agents had a lower linking cost or a higher value. Linking costs and values are different across players. The payoff of a player increases with the value of his partners but decreases with distance and cost.

They designed different treatments. The baseline treatment assumes homogeneity of players. In the other treatments, the experimenters introduced one "low-cost" player or one "high-value" player (or both) in a group of "normal" players to see their contributions. Players had a complete information. Agents knew their own type (normal, high-value or low-cots) and the types of the players in their group.

We do not detail their theoretical results but globally the strict Nash networks are star networks with the "special agent" in the center or the empty network. The next results concerns efficiency. A network is efficient if it maximizes the sum of individual payoffs in the network. For the treatments with a high-value player, the efficient network is the star (CSS or PSS) with the high-value player in the center. Finally for the treatments with a low-cost player, the CSS with the low-cost agent in the center is the efficient network. In summary, the efficient network is a star network for every treatment.

However the star is asymmetric and can lead to uneven payoffs. If players are inequality averse, they can feel envious (they have a lower payoff than the others) or guilty (they have a higher payoff than the others). The authors checked the impact of "natural" levels of inequality. They introduced concerns for fairness in the utility function. They found that theoretically strict Nash networks remain strict Nash even with a concern for fairness.

Here is the description of their experimental procedure and results. The experiment lasted for 30 rounds and groups of players remained the same during the whole experiment (*partner matching*). The results of the experiment are the following. The empty and complete networks never emerged. Their main finding is that the presence of a high-value agent facilitates the formation of equilibrium and star networks. Indeed from 40 to 50% of networks are strict Nash when there is a high-value agent. They are very important for coordination. The presence of a low-cost player did not have the same effect. A detailed analysis shows that stars did not emerge a lot at the beginning of the experiment. However they became increasingly prominent over time. The repetition of the process and the learning effect are important to see the emergence of stars. Globally stars prevailed in treatments with a high-value agent in the center.

Finally a last point is to understand the reasons of the difference between treatments with a high-value agent and with a low-cost agent. The problem with the treatment with a low-cost agent is that this player pays for all the links and feels envious against the others. He can decide to remove all his links because of inequality aversion. However a high-value agent is the center of the PSS at the equilibrium. Peripheral players pay for the links. He does not have incentives to deviate and peripheral players would hurt themselves more by removing a link.

The three issues (asymmetry, efficiency and fairness) described in these papers may play a great role for our experiment. We will probably have experimental results that are not in accordance with our theoretical results. We develop this part later in the paper.

These three papers inspired our protocol. The experiment of Goeree et al. is close to our experimental protocol. However even if strict Nash networks emerge in some situations with a high-value player, the presence of a low-cost player does not facilitate the emergence of equilibrium networks. We add some features in order to solve this issue. We have two major differences with their experimental procedure. The first difference is that link formation is sequential. It brings a new perspective where the market leader (that can be compared with the high-value or low-cost agent in their model) has another advantage: he plays first. It will probably change the emergence of networks. We think that sequentiality will bring more stability. Players can take the time to analyse the actions of the preceding players. They take their decisions in accordance with the actions of the preceding players and with the potential actions of the following players. The second difference is that values and costs are assigned based on players' performance elicited in a previous stage and are not given by the experimenter. We expect that this will reinforce the role of leader, because players earned their position. The aim is testing whether in our setting high-value and low-cost players facilitate the coordination of the network or not.

3 Theoretical model

3.1 Notation

Let G be the set of possible networks. The network $g \in G$ is a pair (N, L) where N is the set of players, with i and j two typical players, and L represents the set of links. In our model we limit the number of players, n, at four. It is quite a small number but there are many potential networks (stars, line, wheel, empty, complete...). A link between player i and player j is denoted by ij. With a slight abuse of notation we write $g_{ij} = 1$ when players i and j has formed a link in g and we write $g_{ij} = 0$ when players i and j has not formed a link in g. If two players i and j have formed a link in g, then i and j are *adjacent* players. We say that they are *directly connected*. A link between two players i and j is identified with the existence of a pairwise relation between them. A *path* between player i and j is a sequence of links $i_0i_1, i_1i_2 \ldots, i_{m-1}, i_m$ where $i_ki_{k+1} \in g$ for $k \in \{0, \ldots, m-1\}$, and $i = i_0$, $j = i_m$. In the following, we say that two players are *indirectly connected* if there exists a path between them but they are not adjacent. In this case these players have an *indirect link*.

The linking actions of a player k are denoted by $g(k)$. As we have four players, each player can decide to form a link (or not) with the three other players, i.e. $g(i) = \{0, 1\}^3$ for all $i \in N$. For player 1, we have $g(1) = (g_{12}, g_{13}, g_{14})$ (similar definition for the other players). Let $g(-i)$ be the actions taken by players other than i. For instance, $g(-1) = \{g(2), g(3), g(4)\}$. With a slight abuse of notation we write $g = \{g(i), g(-i)\}$ for all $i \in N$. Finally, let $g_k^- = \{g(1), ..., g(k-1)\}$ and $g_k^+ = \{g(k+1), ..., g(4)\}$ represent respectively the actions of players that play before and after k.

Let $N_i(g)$ represent the players observed by i and $N_i^d(g)$ be the set of links created by i; $\mu_i(g)$ and $\mu_i^d(g)$ are their respective cardinal numbers. A network $g' \subset g$ is a *component* of g if there is a path between i and j in g' and if for all i in g' and k in g, g_{ik} implies that k also belongs to g'.

Player i needs to pay a connection cost c_i. This cost of linking can be different across players. The costs of linking are a n-tuple $\mathbf{c} = (c_1, ...c_n)$.

Players have also heterogeneous values. The values of players are a n-tuple $\mathbf{V} = (V_1, ...V_n)$ where V_i is the value of player i. In our model information can flow without being altered. In the language of BG we say that there is no *decay*. It means that whatever the distance between two players, information will keep the same value.

As two links can be created between each pair of players and that we have four players, we have 2^{12} potential networks. We define formally some networks that will be useful in the following of the paper.

The *empty network* is the network g in which there is no link at all. On the contrary the *complete network* is a network g when there is a link between all pairs of players. A *star network* is a network g where there is a player, say player i who is adjacent with all other players while players $j \neq i$ are adjacent only with i. We say that i is the *central* player of the star and players $j \neq i$ are *peripheral* players. As our model is directed, we have three types of star networks. In the *periphery sponsored star* (PSS), all peripheral players pay for a link with the central player. On the contrary, the central player is the initiator of all the links with the peripheral players in the *center-sponsored star* (CSS). Finally in the *mixed star* both the central player and peripheral players pay for links. The *wheel* is a network where each player forms only one link so that the architecture of the network looks like a circle. Finally g is *connected* if it has a unique component. A network g is *minimally connected* if the network is connected and there is no cycle in the network (unique path between all pairs of players).

3.2 Structure of the game

We build two treatments (for our model and our experiment).

Cost-treatment

The first treatment is called the cost-treatment (henceforth CT). In this treatment costs are heterogeneous (different c_i for all $i \in N$) but values are homogeneous ($V_i = V$ for all $i \in N$). This treatment depicts situations where some players can create links more easily than others but they have the same value. An agent has a better capacity to form links and to coordinate the network. For instance a firm can have better communication skills than the others. It can collaborate with other firms more easily. Another example is the case of friendship. Some children or adolescents are more sociable. They can create links with other individuals more easily. It does not involve a big effort. They are more popular and are the center of a friendship network.

Value-treatment

The second treatment is called the value-treatment (henceforth VT). In this treatment values are heterogeneous (different V_i for all $i \in N$) but costs are homogeneous ($c_i = c$ for all $i \in N$). It describes situations with more or less valuable agents. Again with the example of firms, some collaborations are more valuable than others. Co-authorship among scientists is another example. Authors have different "values". A PhD student prefers to be linked with a reputed author. His benefit will probably be higher. This student accepts to pay the linking cost to work with this author. The reputed author also benefits from this connection but to a lesser extent.

Sequentiality

A main difference with the major part of the literature on networks is that we chose a sequential setting. It means that players will play one after another. In our model one player is the "market leader" and the others are "followers". The market leader plays first and the followers play after him in a sequential manner. After the market leader made his decisions, the other players can also form links one after another. Let $\rho = i, j, k, l$ be the rule of order. For the CT we rank players according to their cost. Formally, $\rho_i < \rho_j$ if and only if $c_i < c_j$. The lowest-cost player is the market leader and plays first. We assume that a player who has a lower cost of linking can form links more easily. After the first player took his decisions, the second lowest-cost player can play, then the third player and finally the fourth can play (highest-cost player of the group). For the VT we rank players according to their value. In this situation $\rho_i < \rho_j$ if and only if $V_i > V_j$. The highest-value player is the market leader and plays first. We assume that a player with a higher value deserves the position of leader. When the first player has finished to play, the second highest-value player can play, then the third plays and finally the fourth can play (lowest-value player of the group).

3.3 Profit function

The profit of a player i is positively related with the value of his direct and indirect parters and is negatively related with the costs of the links he creates. For the CT, the profit of i is:

$$\pi_i(g) = \mu_i(g)V - \mu_i^d(g)c_i \qquad (1)$$

For the VT, the profit of i is:

$$\pi_i(g) = \mu_i(g)V_j - \mu_i^d(g)c \tag{2}$$

3.4 Equilibrium notion and efficiency

The Nash equilibrium ignores the sequential structure of the game [25]. We need the notion of *subgame perfect equilibrium* where players reassess their plans as play proceeds. The SPE that we present is adapted from Bialas [4]. We need some notations to define our concept.

Let G be the set of networks. As we said earlier there are 2^{12} potential networks. In the network g we write $g_{ij} = 1$ with $i \neq j$ if players i and j have formed a link in g. On the contrary, if they are not connected we write $g_{ij} = 0$. Players make their decision one after another. Let $\rho(g) = 1, 2, 3, 4$ be the rule of order where $\rho = 1$ for the market leader of our game. Each player k can take decisions to form links. The actions of player k are denoted by $g(k)$. Every player decides to form (or not) a link with the three other players. Consequently, $g(1) = (g_{12}, g_{13}, g_{14})$, $g(2) = (g_{21}, g_{23}, g_{24})$, $g(3) = (g_{31}, g_{32}, g_{34})$ and $g(4) = (g_{41}, g_{42}, g_{43})$. To sum up each player has to take three decisions: $g(i) = \{1, 0\}^3$ for all $i \in N$.

We define $g(-1) = \{g(2), g(3), g(4)\}$ as the actions taken by players other than 1. We have a similar definition for player 2, 3 and 4. With a slight abuse of notation, we write $g = \{g(i), g(-i)\}$ with $i \in N$. Finally, let $g_k^- = \{g(1), ..., g(k-1)\}$ and $g_k^+ = \{g(k+1), ..., g(4)\}$ represent respectively the actions of players that play before and after player k. The first term is fixed as these players played before but the second term varies. Player k knows the actions of the preceding players but not the actions of the following players.

Here is the intuition of the equilibrium notion. Each player k wants to maximize his profit denoted by $\pi_k(g)$. However individual profits depend on the actions chosen by the other players. Each player needs to take into account the strategies of the preceding and the following players. Formally, player k has to solve:

$$\max\{\pi_k(g) : (g(k)|g(-k))\}$$
$$st.g \in G$$

The first step to define our equilibrium is to study the case of the last player (fourth player in our case). We begin with the strategy of the last player and then we will study the strategies of the preceding players. It is the principle of *backward induction*.

The strategy of the last player (player 4 in our setting) is the simplest. He has to maximize his profit according to the actions of the preceding players (this set is already fixed). As they are no other players after him he does not have to take into account what the following players may decide. His decision is the last decision of the game and determines all the payoffs. Formally, player 4 has to solve:

$$\max\{\pi_4(g) : (g(4)|g_4^-)\}$$

The strategies of player 1, 2 and 3 (g_4^-) are fixed. Only $g(4)$ can vary.

The actions available for player 4, where g_4^-, is fixed are represented by this set:

$$\overline{G}_{n-1} = \{g \in G : \pi_4(g) = \max\{\pi_4(g) : (g(4)|g(-4))\}\}$$

In our setting, due to transitivity, a player may be indifferent between two actions. Concretely if two players i and j form a component and that a third player l wants to create a link with this component, he is indifferent between creating a link with i or j. In order to have uniqueness of equilibrium we define a tie-breaking rule to define the action chosen if there is indifference between two actions. For that, we look at the *lexicographical order*. To define it quickly, let suppose that we have two pairs $a = (a_1, a_2)$ and $b = (b_1, b_2)$. The set a is lexicographically preferred to b if $a_1 > b_1$ or if $a_1 = b_1$ and $a_2 > b_2$. In our setting if player k has two possible sets of actions and that he is indifferent between the two sets of actions, we choose the lexicographically greater set. Let suppose that we have $g(k) = (1, 1, 0)$ and $g'(k) = (1, 0, 1)$ two sets of actions that leads to the same profit, i.e. $\pi_k(g) = \pi_k(g')$. We say that g is lexicographically preferred to g', i.e. $g \succ g'$. Player k forms two links with the players that play earlier in the game.

Formally the restricted set of feasible actions for the last player is:

$$G_{n-1} = \{g \in \overline{G}_{n-1} : \pi_4(g) = \pi_4(g') \Longrightarrow g(4) \succ g'(4), \forall g' \in \overline{G}_{n-1}\}$$

For the player that plays before the last player (player 3 in our setting), the maximisation problem is the following:

$$\max\{\pi_{n-1}(g) : (g(n-1)|g^-_{n-1})\}$$
$$st.g \in G_{n-1}$$

The third player knows the actions taken by the player 1 and 2. This is fixed. However player 4 has not already chosen his set of actions. When player 3 takes his decision it will define the actions that player 4 should take to maximize his profit.

Then we generalize the problem at the k-level. The set of potential actions are given by:

$$\overline{G}_k = \{\overline{g} \in G_{k+1} : \pi_{k+1}(\overline{g}) = \max\{\pi_{k+1}(g) : (g(k+1)|g^-_{k+1})\}\}$$

We restrict the set of feasible actions with our tie-breaking rule:

$$G_k = \{g \in \overline{G}_k : \pi_{k+1}(g) = \pi_{k+1}(g') \Longrightarrow g(k+1) \succ g'(k+1), \forall g' \in \overline{G}_k\}$$

The maximisation problem of a player k can be written:

$$\max\{\pi_k(g) : (g(k)|g^-_k)\}$$
$$st.g \in G_k$$

We present an example to illustrate the definition of a subgame perfect equilibrium. We assume that $\min c_i < V < \max c_i$ (CT treatment). More precisely we have $c_1 < c_2 < c_3 < V < c_4$ and $3V > c_4$. Like in our definition we begin with the last player. Player 4 wants to maximize his profit. He has an information: the actions of player 1,2 and 3. He has to solve:

$$\max\{\pi_4(g) : (g(4)|g^-_4)\}$$

He has two potential actions. His choice depends on the actions of the preceding players. If he faces the empty network (player 1,2 and 3 did not form any link), he remains passive because otherwise he would loose money. However if he faces a three-player component (player 1,2 and 3 are linked) he has to create one link to be connected to this component. In this second case, he has a strictly positive payoff. Because of perfect transitivity, he is indifferent between creating a link with 1, 2 or 3. Our tie breaking rule stipulates that if a player is indifferent, he decides to link with the player that moves first in the sequential game. In our case, player 4 creates a link with player 1 if he faces a three-player component.

Now, let study the feasible actions of player 3. Player 3 knows the action of 1 and 2 as they played earlier in the game. His action will orientate the action that player 4 has to take to maximize his payoff. The maximisation problem of player 3 is:

$$\max\{\pi_3(g) : (g(3)|g_3^-)\}$$
$$st.g \in G_{n-1}$$

Player 3 can create 0,1,2 or 3 links, i.e. $g(3) = (g_{31}, g_{32}, g_{34})$. If he faces the empty network player 3 has to create links to have a strictly positive payoff. However he does not create more than two links. Indeed he knows that the action of creating two links will "oblige" the fourth player to create one link as player 4 will face a three-player component. Player 3 forms two links, one with player 1 and another one with player 2 because of the tie-breaking rule. Player 2 wants to maximize his profit but knows that his action will impact on the decisions of player 3 and 4. If he does not create links, player 3 and 4 will have to form links to maximize their payoff. We make the same reasoning for player 1.

To sum up, to reach the SPE, player 4 has to create a link with player 1 and player 3 has to create two links, one with player 1 and another with player 2. Formally, $g(1) = g(2) = \emptyset, g(3) = (1, 1, 0)$ and $g(4) = (1, 0, 0)$.

Our tie-breaking rule allows us to have a unique equilibrium. However we can easily define the other equilibrium. For our example, all the networks where player 4 creates one link and player 3 creates two links are SPE.

3.5 Results

We can divide the results into two parts. First we provide the overall results of the model. Second we focus on few interesting and problematic situations by parametrizing the variables.

3.5.1 General results

Players can create links and benefit from the partners they are linked with. If they are the initiator of the link they have to pay a connection cost. The principle of *transitivity* is very important in our model. Players benefit from their direct and indirect links without decay (benefits do not depend on distance). Benefits flow via paths and every player that is part of the component can benefit from these benefits because information flows both ways and without decay. At the maximum players should create three links in the network that involve the four players in a unique component. As we have transitivity, creating more than three links is useless. As a consequence remaining passive can be a good strategy if the other players create a connection with this player.

Remark 1 *Let the payoff be given by* (1) *or* (2)*. A network should always have no more than three links if it creates a unique component of four players.*

The network formation process is dependent on the order. The position of players in the network mainly defines their strategies. The last player is in a very disadvantageous position in many situations in both models if the group wants to reach the equilibrium. As the previous players know that this player will have no choice but creating links if he wants a positive profit, they can remain passive. Consequently the SPE is often the CSS with the last player in the center. In the major part of the situations the last player (highest-cost or lowest-value player) has to sacrifice himself for the others. To reach the equilibrium, the last player has to form as many links as he can. Formally, if player i is the last player of the game and $\mu_i(g)V_j - \mu_i^d(g)c_i > 0$ he has to create some links. If the last player cannot create all the links, the penultimate player has to create the rest of the links to involve the four players in the network. Again, if the links created by the two players are not enough to involve the four players, the second player will create the links. Finally, the first player can create links if the others cannot create all the links necessary to reach the equilibrium.

Remark 2 *Due to backward induction, the last player is the central player of the CSS in numerous situations of both treatments. Consequently, he has to incur all the linking costs.*

About efficiency, for the CT the efficient network is always the CSS with the first player (lowest-cost player) as central player if $V > c_1$. However, if $3V < 3c_1$ the first player cannot have a positive payoff if he creates all the links. In this case, the first player creates as many links as he can (until his payoff is just above zero) and player 2 creates the necessary links to involve the four players in the network. For the VT, every minimally connected network is efficient. The sole condition is to have three links in the network that connect the four players.

Proposition 1 *The efficient network for the CT is the CSS with the first player (say player 1) as central player if* $V > c_1$*. For the VT every minimally connected network is efficient if* $3V_j > c$*.*

Now let just define the conditions to have the empty network for the CT and the VT. For the CT, if $V < c_i$ for all i and $\mu_i(g)V < \mu_i^d(g)c_i$ for all $i \neq j$ the equilibrium network is the empty network. Similarly, for the VT, if $V_i < c$ for all i and $\mu_i(g)V_j < \mu_i^d(g)c$ for all $i \neq j$ the equilibrium network is the empty network. Now let study some specific situations by parametrizing the variables.

3.5.2 Specific situations

We begin with the analysis of particular situations with the CT.

Cost-treatment

In this case, we have heterogeneity in costs but homogeneous values, i.e. $V_i = V$. We rank the players according to their costs. For more simplicity we assume that $c_1 < c_2 < c_3 < c_4$. The lowest-cost player plays first and is the market leader. We present four situations.

First situation

The first situation assumes a high homogeneous value compared to the costs of linking. More precisely we assume $V > \max c_i$. Every player can create all the links.

Second situation

Contrary to the first case, here the value of players is very low. More precisely, we assume $V < c_i < 2V$ for all i. No player can create all the links but if they coordinate, players can form a network.

Third situation

In this case the value V is intermediate, i.e. $\min c_i < V < \max c_i$. More precisely we choose $c_1 < c_2 < c_3 < V < c_4$ and we add a condition: $3V > c_4$. Player 1, 2 and 3 can create all the links but player 4 can also create one link if he is connected with a three-player component.

Fourth situation

In this situation, V is again intermediate. More precisely we have $c_1 < V < c_2 < c_3 < c_4 < 2V$. Player 1 is the only one who can create links individually without depending on the others but the others can create one link if they are connected with a two-player component.

The subgame perfect equilibrium of each situation is presented in the following proposition.

Proposition 2 *Let the payoff be given by* (1).

1. *If* $V > \max c_i$*, the equilibrium network is the CSS with the highest-cost player as central player.*

2. *If* $V < c_i < 2V$ *for all* i *the equilibrium network is a network where the three last players (with the highest costs) create each one link with the first player and the low-cost player remains passive. It is a PSS with the first player as central player.*

3. *If* $c_1 < c_2 < c_3 < V < c_4$ *and* $3V > c_4$*, to reach the equilibrium network player 4 must create a link with player 1 and player 3 must create a link with player 1 and 2.*

4. *If* $c_1 < V < c_2 < c_3 < c_4 < 2V$ *the equilibrium network is a network where the three last players (with the highest costs) create each one link with the first player and the low-cost player remains passive. It is a PSS with the first player as central player.*

Here is the intuition of the proposition for the first situation. Every player is able to connect to everyone in the sense that they all have a cost of linking lower than the value of others. The main component that determines the SPE is the order. Let assume that we have $c_1 < c_2 < c_3 < c_4 < V$, the best response of player 4 to the empty network is to be connected to the rest of the players. Player's 3 best response to the empty network is also to create links with all the players. But he knows that it is also the best response of 4 when the network is empty. We apply the same reasoning for player 2 and 1. Consequently, because of the order and backward induction, the last player (player with the highest cost) has to create all the links. This solution is counter-intuitive as the highest-cost player bears all the linking costs.

In the second situation, contrary to the first case, the value of players is very low. A network can still be formed thanks to transitivity but players need to coordinate because one player cannot create all the links. Indeed if one player is indirectly connected with another player, he benefits from this connection without paying for it. No player can create a link without being connected to a component of at least two players and no player can create two links. The best response of each player to the empty network is to create no link. However the best response of player 4 if it exists a component of two players (a link has already been created between two players) is to connect with this component. Player 3 has the same best response if he faces a component of two players. But he knows that player 4 has the same best response. Each of them can create one link at the maximum. Player 2 has the same best response and knows that 3 and 4 will create a link to be connected to the component. Player 1 knows that the best responses of the followers is either the empty network (in this case player 1 cannot create links) or that each of them will create one link. In the second case his best response is also to remain passive because the followers will do the necessary connections and include him in the network. Each of them must create one link to create a component of four players and allow the diffusion of the benefits thanks to transitivity. Each of the three last players has to form a link with another player. They are indifferent concerning their partner. Our tie-breaking rule suggests that if a player is indifferent between two links, he chooses the partner that plays the earliest in the game. Here player 2, 3 and 4 create one link with player 1. The architecture of the SPE is a PSS with player 1 as central player.

We have the same reasoning for the third situation where V is intermediate, i.e. $\min c_i < V < \max c_i$. More precisely we choose $c_1 < c_2 < c_3 < V < c_4$ and we add a condition: $3V > c_4$. Even if player 4 has a cost higher than V, he can still create one link thanks to transitivity. Indeed we assume that if he is linked to a component of 3 players and that he only creates one link with this component, he has still a positive profit. On the contrary if $3V < c_4$ player 4 cannot do anything and is only dependent on the others. To come back to our case, player 4 wants to be connected to a component of 3 players. However he can only create one link. The best response of player 4 when he faces the empty network is to create no link, otherwise he would loose money. But if he can reach a component of three players his best response is to create a link. Player 3 can create as many links as he wants but he knows that the best response of player 4 when he faces a three-player component is to create a link. Player 3 knows that if he is part of a three-player component, player 4 will link with him (directly or indirectly). The best response of player 3 to the empty network is to create two links. Player 1 and 2 have the same reasoning but know that player 3 will form links if he faces the empty network. Player 4's best response is to create one link with player 1 (because of the tie-breaking rule) and player 3's best response is to create a link with player 1 and 2 (because of the tie-breaking rule). Player 1 and 2 attract the benefits of linking without paying for it.

Finally for the fourth situation V is again intermediate. In the case where $c_1 < V < c_2 < c_3 < c_4 < 2V$, player 1 is the only one who can create links individually without depending on the others. Player 2, 3 and 4 can create one link if they are connected to a component of at least two players. The reasoning is the same than in the previous case. The best response of player 2, 3 and 4 to the empty network is to create zero link. For player 1 the best response to the empty network would be to create three links. However the best response of player 4 if he faces a component of at least two players is to create a link with this component. Player 2 and 3 face the same situation. If we follow our tie-breaking rule, the SPE is a PSS with the first player as central player.

These results seem counter-intuitive as the player with the lowest cost always remain passive even in situations where he could create all the links. This is due to the order of turn and backward induction. We can say that in the four situations, the last player is in a disadvantageous position.

Now let study efficiency. The efficient network of each situation is described in the following proposition.

Proposition 3 *Let the payoff be given by* (1).

1. *If* $V > \max c_i$*, the efficient network is the CSS with the lowest-cost player as central player.*

2. *If* $V < c_i < 2V$ *for all* i *the efficient network is a network where the first three players (with the lowest costs) create each one link and the highest-cost player remains passive.*

3. *If* $c_1 < c_2 < c_3 < V < c_4$ *and* $3V > c_4$*, the efficient network is the CSS with the low-cost player as central player.*

4. *If* $c_1 < V < c_2 < c_3 < c_4 < 2V$ *the efficient network is the CSS with the low-cost player as central player.*

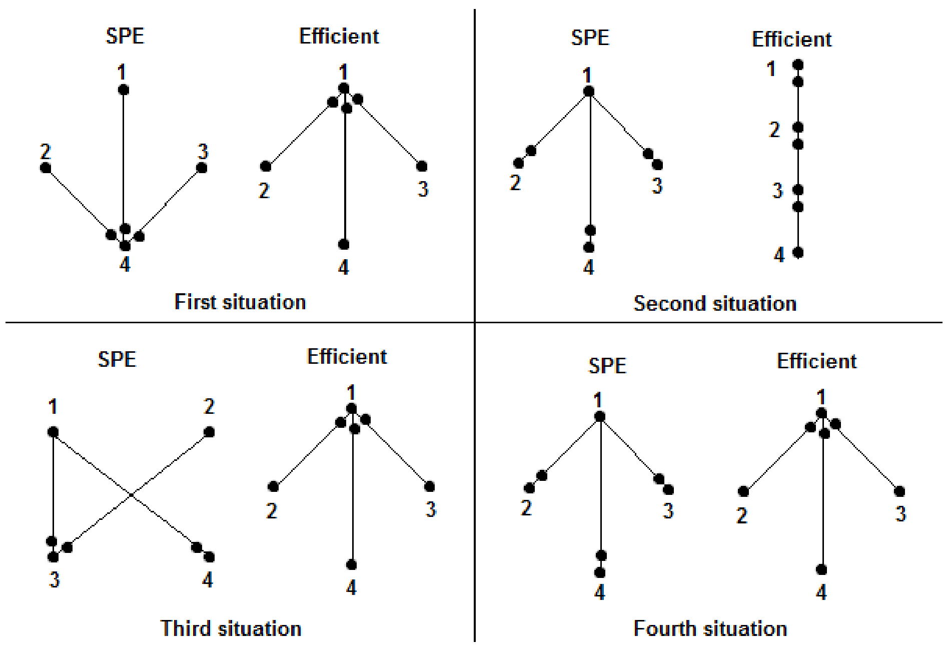
We give the intuition of these results. In the CT, it is less expansive at the aggregate level to form the network if the low-cost player incurs all the costs. That is why the first player is the center of the CSS in three situations.

If we compare with the SPE results, we have non coincidence between the equilibrium and the efficient network.

For the second situation, the first player cannot create all the links because of the low values. They can create one link at the maximum. To minimize the aggregate cost of linking, the first three players (who have a lower cost of linking than player 4) has to create one link. The network can have many different architectures: PSS, line, mixed-star, etc. as long as the first three players create each one link.

We do not have coincidence of equilibrium and efficiency in this treatment. In the CT we have coincidence between the equilibrium and efficiency if and only if the linking cost of the leader is very low and the linking cost of the other players is extremely high. In this case the equilibrium and efficient network is the CSS with player 1 as central player.

Here is a graphic representation of the equilibrium and efficient networks.



When there is a filled circle on a link adjacent to one player, it means that this player created the link. For the SPE, we draw the equilibrium found when we apply the tie-breaking rule. Other networks can be SPE. About efficiency, in the second situation, the efficient network can have different architectures: PSS, mixed-star, etc. We just provide one possible efficient network to give an example.

Value-treatment

In this case we have heterogeneity in values but homogeneous costs of linking, i.e. $c_i = c$. We rank players according to their values. The highest-value player is the market leader. The main finding is that the results are similar with the cost-treatment. The reasoning is quite the same and the order is still the prominent determinant. Let assume $V_1 > V_2 > V_3 > V_4$ for more simplicity.

First situation

We analyse the case where the cost of linking is very low compared to the values of agents. More precisely we have $\min V_i > c$. Every player can create all the links.

Second situation

Now we analyse the opposite situation where the cost of linking is very high compared to the values of players. More precisely we set $c > V_i$ for all i and $3V_j > c$ for all $j \neq i$. At the maximum a player can create one link if players coordinate.

Third situation

As in the second situation, we have a high cost of linking but the gap between the values and this cost is smaller. In this case we have $c > V_i$ for all i and $3V_j > 2c$ for all j. It means that each player can create two links (not only one) if they coordinate.

Fourth situation

Finally the last situation involves an intermediate cost. Let $\max V_i > c > \min V_i$ and $3V_j > 2c$.

The subgame perfect equilibrium of each above situation is given in the following proposition.

Proposition 4 *Let the payoff be given by* (2).

1. *If* $\min V_i > c$ *the equilibrium network is the CSS with the lowest-value player as central player.*

2. *If* $c > V_i$ *for all* i *and that* $3V_j > c$ *for all* $j \neq i$ *the equilibrium network is a network where the three lowest-value players create each one link with the first player. The SPE is a PSS with the first player as central player.*

3. *If* $c > V_i$ *for all* i *and that* $3V_j > 2c$ *for all* j *the equilibrium network is a network where player 4 creates two links (with player 1 and 2) and player 3 creates one link with the first player.*

4. *If* $\max V_i > c > \min V_i$ *for all* i *and that* $3V_j > 2c$ *for all* j *the equilibrium network is a network where player 4 creates two links (with player 1 and 2)and player 3 creates one link.*

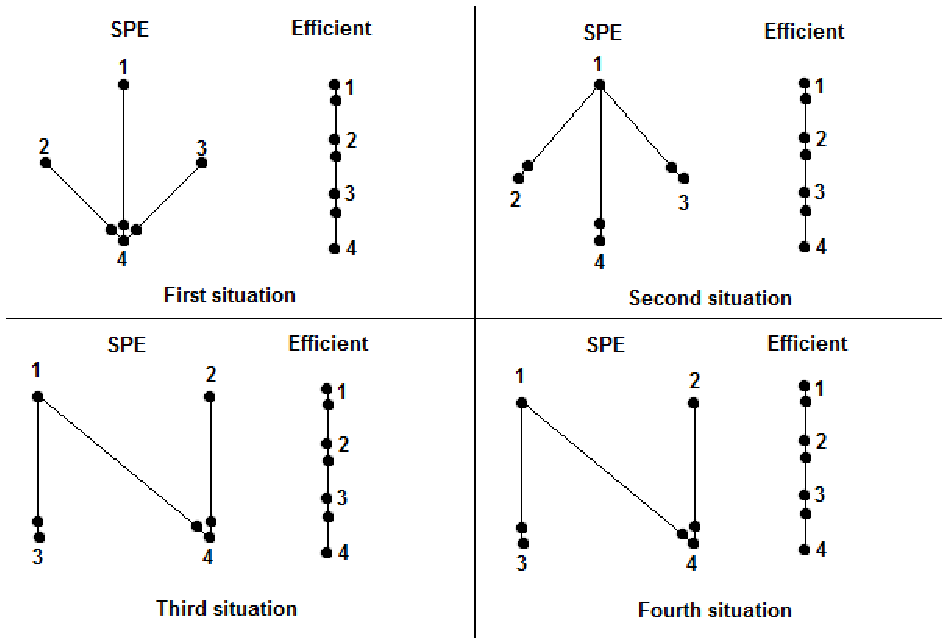
The reasoning is the same than in the cost-treatment. For the first situation, we assume $V_1 > V_2 > V_3 > V_4 > c$. Every player can create all the links and keep a positive payoff. The best response of player 4 when he faces the empty network is to create all the links. Player 1, 2 and 3 could also play this strategy but they have the information that the following players can bear the costs. Player 4 has no choice but creating all the links. The order is still very decisive. With transitivity, it is not necessary to have the high-value player as central player as soon as he is part of the network.

With the second situation we can say that players cannot create links separately. But if players coordinate they can form a network thanks to transitivity. Player 1 knows that players 2, 3 and 4 can each create a link they face a three-player component. Player 1 remains passive. This minimally connected network involving the four players can take many architectures: line, PSS, mixed-star, etc. If we apply our tie-breaking rule, the SPE is the PSS network with the first player as central player. We do not develop the reasoning for the last two situations as it is exactly the same principle than in the CT. Now let study the efficient network of each situation.

Proposition 5 *Let the payoff be given by* (2). *In every situation, the efficient network is a minimally connected network. The sole condition is to create three links involving the four players in a unique component.*

It means that whoever creates links, as soon as the 4 players are connected, the network is efficient. We find this result because the cost is the same for all connections, so it does not change the aggregate payoff. In the VT the equilibrium and the efficient network can coincide more frequently.

Here is a graphic representation of the above propositions.



For the SPE, we draw the equilibrium found when we apply the tie-breaking rule. Other networks can be SPE. For efficiency, networks can have different architectures. We just provide an example of solution.

To sum up, both treatments have the same reasoning and the order seems to be the prominent determinant. We only have one difference between the two treatments. The difference is that we can have coincidence between the equilibrium and the efficient network in the VT while it is almost impossible in the CT.

To conclude, we can say that star networks and minimally connected networks are prominent in our model. Globally the same architectures emerge than in the connection model of Bala and Goyal, but we have more stability thanks to heterogeneity. About efficiency a natural candidate emerges to take the position of central player in the CT: the low-cost player. But in the VT every minimally connected network is efficient. The high-value player is not always the center of the star contrary to the paper of Goeree et al [19]. This is due to perfect transitivity (benefits do not decrease with distance in our model). A final word about this model is to say that playing first is often an advantage and playing last can be a very bad position in the network. The market leader has often the highest payoffs.

4 The experiment

The formation of networks with heterogeneous players is hard to study in the field. It is difficult to measure the values and linking costs of agents. In contrast, the laboratory enables us to control all the variables that impact the behaviour of players. The three papers ([3], [13] and [19]) presented in our experimental literature inspired our experimental procedure.

We have two treatments, an endogenous and an exogenous treatment (baseline treatment). Each treatment consists of a CT and a VT (heterogeneous costs and a homogeneous value or heterogeneous values and a homogeneous cost). In the endogenous treatment, the experiment consists of two parts. In the first part, the values and the costs of linking are revealed by each player thanks to tests. In the second part, players decide in a sequential manner with whom they want to form a link according to these variables.

The baseline treatment only consists of the network formation part. Contrary to the endogenous treatment, we randomly assign characteristics for the players.

4.1 The baseline treatment

We begin the presentation of the experiment with the baseline treatment. Participants are randomly assigned to a group of four players. The purpose is to see which types of networks emerge. Our model gives $2^{12} = 4096$ potential networks. For the sake of simplification, we focus on the specific situations described in the theoretical results.

We assign one specific situation to each group of four players. They play either the CT or the VT (not both) and remains in the same situation (for example the second situation of the CT) during the whole experiment. The design of our experiment is based on a between-subject design because we want to avoid that a treatment impacts on the other treatment and because the order of exposure can change the behaviour of participants.[1] When groups are formed and that each group is assigned to a specific situation, we parametrize the costs and values and assign them randomly within the

[1] Charness, Gneezy and Kuhn [7] explained the differences between these two designs and their advantages and disadvantages.

group. The four situations for the CT are described in the first column of the following table.

Cost-treatment	Subgame Perfect Equilibrium	Efficient network
1st situation: $c_1 = 1$, $c_2 = 2$, $c_3 = 3$, $c_4 = 4$ and $V = 5$	CSS (player 4 as central player) $\Pi_1 = \Pi_2 = \Pi_3 = 15$ and $\Pi_4 = 3$ and $\Pi_T = 48$	CSS (player 1 as central player) $\Pi_1 = 12$, $\Pi_2 = \Pi_3 = \Pi_4 = 15$ and $\Pi_T = 57$
2nd situation: $c_1 = 16$, $c_2 = 17$, $c_3 = 18$, $c_4 = 19$ and $V = 10$	Minimally connected (2-3-4) $\Pi_1 = 30$, $\Pi_2 = 13$, $\Pi_3 = 12$, $\Pi_4 = 11$ and $\Pi_T = 66$	Minimally connected (1-2-3) $\Pi_1 = 14$, $\Pi_2 = 13$, $\Pi_3 = 12$, $\Pi_4 = 30$ and $\Pi_T = 69$
3rd situation: $c_1 = 1$, $c_2 = 2$, $c_3 = 3$, $c_4 = 10$ and $V = 4$	Minimally connected (player 3: 2 links, player 4: 1 link) $\Pi_1 = \Pi_2 = 12$, $\Pi_3 = 6$, $\Pi_4 = 2$ and $\Pi_T = 32$	CSS (player 1 as central player) $\Pi_1 = 9$, $\Pi_2 = \Pi_3 = \Pi_4 = 12$ and $\Pi_T = 45$
4th situation: $c_1 = 3$, $c_2 = 5$, $c_3 = 6$, $c_4 = 7$ and $V = 4$	Minimally connected (2-3-4) $\Pi_1 = 12$, $\Pi_2 = 7$, $\Pi_3 = 6$, $\Pi_4 = 5$ and $\Pi_T = 30$	CSS (player 1 as central player) $\Pi_1 = 3$, $\Pi_2 = \Pi_3 = \Pi_4 = 12$ and $\Pi_T = 39$

For example, the first group of four players is assigned to the first situation. Players in this group have a high value (5) compared to their cost of linking. We assign a cost of 1 to one player, another player has a cost of 2, another has a cost of 3 and the last player has a cost of 4. In each group, players have a complete information. They know their own cost and value and the costs and value of the other group members. It can be a problem because players are aware of the inequalities between them. The gap can be wide and players may fell envious or guilty.

As we said players move one after another and we rank them according to their cost. The leader is the player with the lowest cost ($c_1 = 1$) and plays first. Player 2, 3 and 4 play after him in a sequential manner. The leader can form links with the three other players. He can decide to form $0, 1, 2$ or 3 links. The three players cannot move but are present in the game. In this situation, for player 1 a link is valuable (he earns 5 EMUs[2] by link) but he has to pay a cost of 1 EMU for each link created. When he has taken his decisions of linking, player 2 can play. Player 2 can observe the choices of player 1 and make his decisions. An important point is to see if player 1 created a link with him or not. If player 1 created a link with player 2, player 2 earns 5 EMUs without paying a linking cost, because player 1 is an indirect neighbour. Player 2 can also link with the three players. When player 2 has finished to play, it is the turn of player 3 and then of player 4. They can all

[2]Experimental Monetary Unit

observe the previous decisions of the game. When the fourth player confirms his decisions, the network is formed. No player can modify his actions. The profit for this period is calculated and given to the participants. They know their own profit for this period but not the profit of the three other players. They repeat the network formation process with the same group during 20 periods.

The situations for the VT are depicted in the first column.

Value-treatment	**Subgame Perfect Equilibrium**	**Efficient network**
1st situation: $V_1 = 5$, $V_2 = 4$, $V_3 = 3$, $V_4 = 2$ and $c = 1$	CSS (player 4 as central player) $\Pi_1 = 9$, $\Pi_2 = 10$, $\Pi_3 = 11$, $\Pi_4 = 9$ and $\Pi_T = 39$	All minimally connected $\Pi_T = 39$
2nd situation: $V_1 = 8$, $V_2 = 7$, $V_3 = 6$, $V_4 = 5$ and $c = 15$	Minimally connected (2-3-4) $\Pi_1 = 18$, $\Pi_2 = 4$, $\Pi_3 = 5$, $\Pi_4 = 6$ and $\Pi_T = 33$	All minimally connected $\Pi_T = 33$
3rd situation: $V_1 = 9$, $V_2 = 8$, $V_3 = 7$, $V_4 = 5$ and $c = 10$	Minimally connected (player 4: 2 links, player 3: 1 link) $\Pi_1 = 20$, $\Pi_2 = 21$, $\Pi_3 = 12$, $\Pi_4 = 4$ and $\Pi_T = 57$	All minimally connected $\Pi_T = 57$
4th situation: $V_1 = 10$, $V_2 = 8$, $V_3 = 7$, $V_4 = 6$ and $c = 9$	Minimally connected (player 4: 2 links, player 3: 1 link) $\Pi_1 = 21$, $\Pi_2 = 23$, $\Pi_3 = 15$, $\Pi_4 = 7$ and $\Pi_T = 66$	All minimally connected $\Pi_T = 66$

Like in the CT each situation is assigned to one group of four players. The difference with the CT is that there is a unique cost inside each group but values differ across players. If we take the first situation, player 1 has the highest value. If player 2 forms a link with player 1, he benefits from the value of player 1 (5) but incurs a linking cost equals to 1. We have the same procedure than in the CT. The network formation process is similar than in the CT and also lasts 20 periods.

We use a partner matching protocol (groups remain the same during the experiment) to facilitate the coordination and the learning effect.[3] The learning effect and the coordination is easier if players remain partners. But to avoid that a player in "bad position" (for example the low-value player in the VT) remains in this position during the whole experiment (he has to bear all the linking costs if the group wants to reach the equilibrium) we rematch the groups in the middle of the experiment. There is a probability

[3]Van Leeuwen, Offerman and Schram [30] point out that coordination is even more complicated with a stranger matching.

that by changing groups players will be at a different position in the network.[4] Each group repeats the network formation process for ten periods before we rematch the groups. Players do not know exactly when we will change the groups to avoid the change of behaviour during the last period before the rematching.

Payoff

In the baseline treatment, the profit only depends on the network formed inside the group. The profit of player i in the network game, $\pi_i(g)$, depends on the value of his partners and on his cost of linking. If player i plays the CT, his payoff is the following:

$$\pi_i(g) = \mu_i(g)V - \mu_i^d(g)c_i$$

If player i plays the VT, his payoff is the following:

$$\pi_i(g) = \mu_i(g)V_j - \mu_i^d(g)c$$

We randomly choose one period and reward them with their behaviour in this period (random draw). This method encourages the players to do their best for each period.[5] The second and third columns of each table gives the profit of each player and the aggregate profit if they reach the equilibrium or if they form an efficient network.

4.2 The endogenous treatment

Like in the baseline treatment, we have two treatments. The first treatment (CT) assumes heterogeneity in costs but homogeneous values. The second (VT) assumes heterogeneous values but homogeneous costs. The difference with the baseline treatment is that we add a preliminary stage. The experiment consists of two parts. The first part is the revelation of values and costs by each player thanks to tests. Consequently their characteristics are

[4] The probability of changing ranking depends on the initial ranking of the players. Obviously it does not solve the problem for the "extreme" players. Indeed the lowest-value, highest-value, lowest-cost and highest-cost players are always in the same position in a group.

[5] Alternatively, we could have decided to take the average of the payoffs during the whole experiment. Similarly we could also decide to sum the payoff of each period. The problem is that players who are in "bad positions" have lower profits on average. For example, if the low-value player is constrained by the group to create all the links, his payoff will be smaller and the gap with the other players will increase. We will have high inequalities.

endogenous instead of arbitrary. The second part is the network formation process. This part is exactly the same than in the baseline treatment except that the costs and values are not assigned randomly.

In the major part of the literature, characteristics are exogenous. Endogeneity brings a new perspective, because players are at the origin of their characteristics. The comparison between the exogenous treatment and this treatment allows the check the contribution of endogenous characteristics.

In the first stage, players make some tests to reveal their characteristics. They are rewarded and ranked according to their performance. Concerning the revelation part, participants in the CT only reveal their cost. The value of players is given exogenously. Symmetrically, participants in the VT only reveal their value. The linking cost is given by the experimenter. Like in the exogenous treatment, we use a between-subject design, so every player only plays one treatment (CT or VT).

Puzzle

Participants play this game if they play the value-treatment. The goal of this test is to assess the value, the ability and the creativity of each player. This task does not require any particular technique. Players just need organization, concentration and ability to solve it as fast as they can. There are no initial inequalities between players. There is time constraint and the fastest player receive the highest value. The slowest player to solve the puzzle will have the lowest value.

English test

Participants play this game if they play the cost-treatment. The linking costs of agents are revealed by an English test (the mother tongue of our participants is French). We assume that the capacity in English reveals the communication skill of the player because English allows communication around the world. Words in French will appear one by one on their computer. Participants will have to translate them into English. They can take all the time they want before confirming their answer. There will be twenty words in total from the easiest to the hardest. We relate negatively the performance at this test with the costs of linking. When players answer correctly, their linking cost decreases. The best player will have the lowest cost. The player who gave the fewest number of correct answers will have the highest cost.

If two players have the same score (for the puzzle or the English test), the software randomly chooses the order between the two players. At the end of the task they know their ranking but not how much they earned exactly to avoid *wealth effect*.

After the revelation part, players are randomly assigned to a group of four participants. Their assignment does not depend on their performance. Then, we assign a situation to each group. They play in the same situation during the whole experiment. Inside each group of four players, we rank them according to their performance. In the CT, we assign the lowest cost to the player that gave the maximum number of correct answers (compared with the three other group members). The second best player has the second lowest cost. The fourth player of the group has the highest cost of the situation. Similarly, in the VT, we assign the highest value to the player that solved the puzzle faster (compared with the three other group members). The second best player has the second highest value. The fourth player of the group has the lowest value of the situation. Players know the values and costs of each player in their group (including their own cost and value).

For example, in the first situation of the CT we assign a cost of linking equal to 1 to the player who gave the maximum number of correct answers during the English test (compared with his three group members). The player with the lowest number of correct answers will have a linking cost of 4.

Even if players do not determine exactly their value or cost, their performance is at the origin of their position in the network. The literature shows that endogeneity counters the impact of social motives. Participants have more strategic behaviours if they earned their position. Here the leader is in the first position because he performed better than the three other players of his group.

Payoff

The total payoff of the players in the endogenous treatment depends on two things: their profit during the task (puzzle if VT or English test if CT) and their profit in the network formation game. The profit in the network game, $\pi_i(g)$, depends on the value of their partners and on the costs of linking:

$$\pi_i(g) = \mu_i(g)V_j - \mu_i^d(g)c_i$$

Then we add a more or less high supplement according to their performance in the first part.

Experimentally, it is not possible to test all the specific situations. We need many observations of one situation to analyse the results. It would require many experimental sessions. For the experiment, the most interesting cases are the first situation of the CT and the VT where the cost is low compared to the value of players. Every player can create all the links. The analysis of these two situations will enable us to see the contribution of sequentiality and heterogeneity in the network formation process.

5 Predictions

According to our model, participants should play the SPE. However our behavioural predictions are different. Efficiency and fairness concerns may affect the emergence of equilibrium networks. Moreover the experiment will highlight the role of objective characteristics thanks to the comparison between the endogenous and the exogenous treatment.

5.1 Theoretical predictions

Our model predicts that participants should play the equilibrium. We just remind that in this case the last player incurs the major part of the linking costs. In every situation of both treatments, the last player creates from one to three links. On the contrary, the leader does not have to create any links. This result is counter-intuitive as the leader does not create links.

However players may be guided by social factors. We present our behavioural predictions.

5.2 Social motives

Players make a tradeoff between self-interest and social motives. As Burger and Buskens said "human subjects are more likely to choose equal and efficient networks as the preferred networks" [5].

Efficiency concern This social motive may influence the behaviour of the participants if they prefer a high aggregate payoff. We want to analyse in which condition the efficient network emerges and more particularly which network emerges between the equilibrium and the efficient network. The efficiency concern has been widely witnessed in network experiments. Charness and Rabin [8] made tests to measure the efficiency concern and found that participants do not always choose to maximize their own payoff. Participants play a more efficient network if it does not increase inequalities too much.

As we pointed out earlier the major difference between the VT and the CT is that in the former the equilibrium and efficiency can coincide while it is not possible in the latter (in the situations we chose to study). In the CT the efficient network emerges when the first player creates all the links or the first three players create each one link. In the VT it is far more easily to form an efficient network. Indeed all minimally connected networks are efficient. Three links that connect the four players are sufficient to create an efficient network.

Hypothesis 1 *Efficiency concerns will limit the emergence of equilibrium networks. Many efficient networks should emerge in the VT. However in the CT we do not think that the leader will sacrifice himself to reach the efficiency.*

Groups may try to tend to an efficient network. For instance, player 1 can decide to form one link to be closer to the efficient network but without bearing all the costs.

Inequality aversion Experimentally, many authors witnessed the impact of inequality aversion ([8], [14], [18]). Even if players are self-interested, they prefer equitable outcomes. In some games, players decide to behave fairly while the model predicts unfair behaviour. In our setting it means that players will share the linking cost and not let the last player sacrifice himself.

When there are inequalities of payoffs players can feel envious (if their payoff is lower than the other players) or guilty (if their payoff is greater than the other players). It is the case in the paper of Goeree et al. [19] and the paper of Falk and Kosfeld [13] with the formation of stars that are very unequal. The PSS of Berninghaus et al. [3] leads to less inequality in payoffs because each peripheral player pays for one link. The most extreme case is the experimental paper of Vanin [31] where players were able to talk before the game. The inequality aversion was so strong within each group that sometimes it prevented the emergence of efficient networks. Some groups decided to play the efficient network despite the inequalities but they redistributed the payoffs after the experiment.

In our experiment, the VT treatment is more egalitarian than the CT. In fact, the SPE and the efficient network in the CT are often the CSS where one player bears all the costs of linking (first player if the group plays the efficient network and last player if they play the equilibrium). On the opposite the VT is less unequal. All minimally connected networks are efficient and not obviously unequal. The SPE can also be the CSS but as the fourth player benefits from the value of the others (that are greater than his own value) inequalities are small even if he bears all the linking costs.

Hypothesis 2 *Inequality aversion will limit the emergence of star networks. The problem of inequality aversion is more prominent in the CT.*

5.3 Coordination

Coordination can be very difficult in network formation. Many previous network experiments witnessed this problem.

First we may have coordination problems to form star networks. Falk and Kosfeld [13] found that the connection model predicts well the behaviour of players in the one-way flow model but not with the two-way flow model. One reason (it does not explain everything) is that it is harder to coordinate on the CSS. The emergence of the CSS requires that one player creates all the links and the others have to remain passive. It may be difficult in our experiment that the four players understood that in some situations they should remain passive for example. Berninghaus et al. [3] had less coordination problems with their experiment because they changed the environment. For participants it is easier to coordinate on the PSS than on the CSS. Another reason is symmetry. The CSS is very asymmetric strategically and in terms of payoffs.

Hypothesis 3 *Coordination problems may restrain the number of star networks and more particularly of CSS.*

Second, we may have a coordination problem for situations where value is low or cost is high. In the CT some situations may be difficult for the players. When value is low (4th situation) or very low (2nd situation), players have to work together to have a positive payoff. Similarly in the VT the situations where the cost is high (3rd situation) and very high (2nd situations), players have to coordinate.

5.4 Endogeneity contribution

This experiment has an additional purpose. We want to analyse the role of some players in the network and more particularly the role of the market leader.

Leadership, responsibility and legitimacy
The market leader has an advantage in our game. He has a high value or a low cost compared to the others. Moreover he takes his decisions first. He is going to orientate the network. In the VT the first player has a real advantage as the other players want to be connected with him. If the group wants to form the equilibrium network the first player needs to remain passive. However to create an efficient network, he can decide to form one link (or more). In this case it will be interesting to see if the first player takes his role of market leader and tries to build a network that is efficient and egalitarian or if he is going to remain passive because he knows he is valuable. We may have two different types of market leader in the VT: leaders or free-riders.

In the paper of Figuières, Masclet and Willinger [15] the authors also made a sequential game. Players can make voluntary contributions to a public good in a sequential manner and can (or not) observe the contributions of the previous players. They found that contribution increases when they can observe previous contributions. They made a second observation that is interesting for our experiment: the leadership effect. The first player contributes more to encourage the others to make a larger contribution. This effect vanishes when the game moves toward the last player.

In the CT the first player has less advantage. If the group tends to the equilibrium the market leader remains passive. However if the first player wants the network to be efficient he has to create the maximum of links he can. In this case he sacrifices himself and bears some linking costs. Here the difference between the two situations is even more striking. In the first situation the leader only benefits from the others and in the second situation the others benefit from him. The difference between the two types of market leaders is even stronger. The first player may decide to free-ride or to take his role of leader to coordinate the network to an efficient network by creating all the links.

To sum up, the leader may have two different types of behaviour.

Hypothesis 4 *The leader may decide to be selfish or to be the coordinator of the network. We think that the first player will be more altruistic in the CT to reach an efficient network even if he will bear some linking costs. In the VT we do not think that the first player will create links as he knows that he is the most valuable and that players will link with him.*

Contribution of objective characteristics

One major difference in our procedure compared to the rest of the literature on network formation is that the characteristics of the players are not exogenous nor arbitrary. They depend on their performance in a preliminary task. We think that this change will impact the behaviour of the agents. That is the reason why we created a control treatment with random variables. We want to control the effect of objective characteristics.

The literature ([9], [20] and [26]) witnessed the "earnings effect". The authors show that when a player earned his dominant position, he behaves more strategically than if this position has been randomly assigned. He behaves more strategically because he feels legitimate. He has a moral authority. In our endogenous treatment the leader should be more strategic than in the exogenous treatment as he deserves his position. He will probably remain passive and let the others create the links. Similarly, the fourth player should accept his position more easily than if we ranked him randomly at this place as his performance was worse than his three group members. He accepts to create links even if it is not advantageous.

Hypothesis 5 *We think that the first player will be more strategic in the endogenous treatment than in the exogenous treatment. The fourth player should accept his position in the network more easily and bears some linking costs. Endogeneity should partly counter the impact of social motives.*

5.5 Punishment

In some cases players have the power to punish his group members. We present one example to illustrate it. In the first situation of the cost-treatment, the SPE is the CSS with player 4 as central player and the efficient network is the CSS with player 1 as central player. If the group wants to reach the equilibrium, player 4 will have to bear all the linking costs. It may be possible that player 4, when he faces the empty network, decides to create a link with player 2 and 3 but not with player 1. It would be a sanction for player 1, a signal sent to player 1 to incite him to create links. Even if he looses money for this period (a link with player 1 would increase his payoff) it can be a way to orientate the network to a more efficient network.

5.6 Protect the little brother

Finally, a last behaviour may appear in our experiment. In the CT the last player has a high cost of linking and it may be very costly for him to be part of a component. The other players may feel altruistic toward this player. They can create a link with him to be sure that he will be part of the component without having to pay a high linking cost. It would protect the "little brother". This behaviour may be guided by altruism (warm glow or pure altruism).

6 Conclusion

Network formation is a very wide field with diverse applications. Bala and Goyal show that networks can be unstable. Our goal is to add two properties to their model to solve instability: heterogeneity and sequentiality. Players are heterogeneous in the sense that they are more or less valuable and they have different communication skills (heterogenous values and costs). We built two treatments: a cost-treatment (with heterogeneous costs but a homogeneous value) and a value-treatment (with heterogeneous values and a homogeneous cost). Players can create links unilaterally and benefit from their direct and indirect links. Information flows both ways, so we have a perfect transitivity. As we add a sequentiality property, players will play one after another. The leader (most valuable or lowest cost player) plays first. The profit of a player increases with the number of connections but decreases with the cost of linking. We display the subgame perfect equilibrium of the game. The first major result is that no more than three links should be created if the four players are involved in a unique component because of transitivity. The equilibrium network architecture is often the star network (CSS). The second important result is that the player in the last position is in a disadvantageous situation. Indeed the preceding players can benefit from him by remaining passive. When it is his turn the last player has no choice but creating the maximum of links to have a positive payoff. Finally we found out that the equilibrium network is not always efficient. Stability and efficiency rarely coincides.

We would like to test the model in the laboratory to see if our model predicts well the behaviour of the participants. Many features may have an important impact on the behaviour of participants. Three aspects of the experiment will be interesting to study. The first purpose is to analyse if decentralized decisions lead to an efficient or an equilibrium network. The second goal is to measure the role of social motives. Indeed human subjects are not only guided by their self-interest, they care about the payoff of others. Finally, our experimental protocol has one particular characteristic. Contrary to the major part of the literature, we build an endogenous treatment where we do not give exogenous values and costs. Each player reveals his personal ability and communication skills in two preliminary tests. Thanks to the comparison between the exogenous and the endogenous treatment, we will be able to describe the "earning effect". It means that players behave more strategically if they earned their position than if it is arbitrary.

Further research is needed to have a better understanding of this complex process. A first extension to this research paper is to introduce cheap talk in the procedure. The idea is to give to the first player (leader) the power to communicate with others. The purpose is to facilitate coordination and to analyse the leadership effect. If behaviours of punishment and protection of the "little brother" appear in the experimental results, it would be interesting to deeper analyse these phenomena. Finally, the study of homophily (tendency to link with similar agents) is another interesting extension. Homophily plays a major role in many real life situations. This principle can be applied to marriage, friendship, relationships at work or information diffusion for instance. In our setting it would be interesting to analyse if players prefer to form connections with similar players or not.

References

[1] Venkatesh Bala and Sanjeev Goyal. A noncooperative model of network formation. *Econometrica*, 68(5):1181–1229, 2000.

[2] Paul Belleflamme and Francis Bloch. Market sharing agreements and collusive networks*. *International Economic Review*, 45(2):387–411, 2004.

[3] Siegfried K Berninghaus, Marion Ott, and Bodo Vogt. On networks and

stars: Recent results in network experiments. *Manuscrito. University of Karlsruhe*, 2004.

[4] Wayne F Bialas. Cooperative n-person stackelberg games. In *Proceedings of the 28th IEEE Conference on Decision and Control*, volume 3, pages 2439–2444, 1989.

[5] Martijn J Burger and Vincent Buskens. Social context and network formation: An experimental study. *Social Networks*, 31(1):63–75, 2009.

[6] Antoni Calvo-Armengol and Matthew O Jackson. The effects of social networks on employment and inequality. *The American Economic Review*, 94(3):426–454, 2004.

[7] Gary Charness, Uri Gneezy, and Michael A Kuhn. Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, 81(1):1–8, 2012.

[8] Gary Charness and Matthew Rabin. Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, 117(3):817–869, 2002.

[9] Todd L Cherry, Peter Frykblom, and Jason F Shogren. Hardnose the dictator. *American Economic Review*, pages 1218–1221, 2002.

[10] Dean Corbae and John Duffy. Experiments with network formation. *Games and Economic Behavior*, 64(1):81–120, 2008.

[11] Sergio Currarini, Matthew O Jackson, and Paolo Pin. An economic model of friendship: Homophily, minorities, and segregation. *Econometrica*, 77(4):1003–1045, 2009.

[12] Dirk Engelmann and Martin Strobel. Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *The American Economic Review*, 94(4):857–869, 2004.

[13] Armin Falk and Michael Kosfeld. It's all about connections: Evidence on network formation. Technical report, IZA Discussion paper series, 2003.

[14] Ernst Fehr and Klaus M Schmidt. Fairness, incentives, and contractual choices. *European Economic Review*, 44(4):1057–1068, 2000.

[15] Charles Figuieres, David Masclet, and Marc Willinger. Vanishing leadership and declining reciprocity in a sequential contribution experiment. *Economic Inquiry*, 50(3):567–584, 2012.

[16] Taiji Furusawa and Hideo Konishi. Free trade networks. *Journal of International Economics*, 72(2):310–335, 2007.

[17] Andrea Galeotti, Sanjeev Goyal, and Jurjen Kamphorst. Network formation with heterogeneous players. *Games and Economic Behavior*, 54(2):353–372, 2006.

[18] Anita Gantner, Werner Güth, and Manfred Königstein. Equitable choices in bargaining games with joint production. *Journal of Economic Behavior & Organization*, 46(2):209–225, 2001.

[19] Jacob K Goeree, Arno Riedl, and Aljaž Ule. In search of stars: Network formation among heterogeneous agents. *Games and Economic Behavior*, 67(2):445–466, 2009.

[20] Elizabeth Hoffman and Matthew L Spitzer. Entitlements, rights, and fairness: An experimental examination of subjects' concepts of distributive justice. *The Journal of Legal Studies*, pages 259–297, 1985.

[21] Matthew O Jackson. A survey of network formation models: stability and efficiency. *Group Formation in Economics: Networks, Clubs, and Coalitions*, pages 11–49, 2005.

[22] Matthew O Jackson and Asher Wolinsky. A strategic model of social and economic networks. *Journal of economic theory*, 71(1):44–74, 1996.

[23] Michael Kosfeld. Economic networks in the laboratory: A survey. *Review of Network Economics*, 3(1), 2004.

[24] Mark EJ Newman. Scientific collaboration networks. i. network construction and fundamental results. *Physical review E*, 64(1):016131, 2001.

[25] Martin J Osborne and Ariel Rubinstein. *A course in game theory*. MIT press, 1994.

[26] David Reinstein and Gerhard Riener. House money effects on charitable giving: an experiment. *Working Papre, University of Essex*, 2009.

[27] Arno Riedl. Efficient coordination in weakest-link games. 2011.

[28] Arno Riedl and Aljaz Ule. Exclusion and cooperation in social network experiments. *Unpublished Paper, CREED, University*, 2002.

[29] Heinrich von Stackelberg. theory of the market economy. 1952.

[30] Dennie Van Dolder and Vincent Buskens. Social motives in network formation: An experiment. In *Game Theory for Networks, 2009. GameNets' 09. International Conference on*, pages 593–602. IEEE, 2009.

[31] Paolo Vanin. Network formation in the lab: a pilot experiment. *Universitat Pompeu Fabra*, 2002.

[32] Louis L Wilde. On the use of laboratory experiments in economics. In *Philosophy in Economics*, pages 137–148. Springer, 1981.