UNIVERSITY OF CAMBRIDGE
Faculty of Economics

Institute for
New Economic Thinking

Cambridge Working Papers in Economics

Cambridge-INET Working Paper Series No: 2020/29

Cambridge Working Papers in Economics: 2063

LARGE SCALE EXPERIMENTS ON NETWORKS: A NEW PLATFORM WITH APPLICATIONS

Syngjoo Choi
(Seoul National University)

Sanjeev Goyal
(University of Cambridge)

Frédéric Moisan
(University of Cambridge)

This paper presents a new platform for large scale networks experiments in continuous time. The versatility of the platform is illustrated through three experiments: a game of linking, a linking game with public goods, and a linking game with trading and intermediation. Group size ranges from 8 to 100 subjects.

These experiments reveal that subjects create sparse networks that are almost always highly efficient. In some experiments the networks are centralized and unequal, while in others they are dispersed and equal. These network structures are in line with theoretical predictions, suggesting that continuous time asynchronous choice facilitates a good match between experimental outcomes and theory. The size of the group has powerful effects on individual investments in linking and effort, on network structure, and on the nature of payoff inequality. Researchers should therefore exercise caution in drawing inferences about behaviour in large scale networks based on data from small group experiments.

JEL: C92, D83, D85, Z13.

Large Scale Experiments on Networks
A New Platform with Applications

Syngjoo Choi* Sanjeev Goyal† Frédéric Moisan‡

July 15, 2020

Abstract

This paper presents a new platform for large scale networks experiments in continuous time. The versatility of the platform is illustrated through three experiments: a game of linking, a linking game with public goods, and a linking game with trading and intermediation. Group size ranges from 8 to 100 subjects.

These experiments reveal that subjects create sparse networks that are almost always highly efficient. In some experiments the networks are centralized and unequal, while in others they are dispersed and equal. These network structures are in line with theoretical predictions, suggesting that continuous time asynchronous choice facilitates a good match between experimental outcomes and theory. The size of the group has powerful effects on individual investments in linking and effort, on network structure, and on the nature of payoff inequality. Researchers should therefore exercise caution in drawing inferences about behaviour in large scale networks based on data from small group experiments.

JEL: C92, D83, D85, Z13.

*Department of Economics, Seoul National University. Email: syngjooc@snu.ac.kr

†Faculty of Economics and Christ's College, University of Cambridge. Email: sg472@cam.ac.uk

‡Faculty of Economics, University of Cambridge. Email: fm442@cam.ac.uk

The paper was supported by the E.U.'s Future and Emerging Technology-Open Project IBSEN (contract no. 662725), the Cambridge-INET Institute and the Keynes Fund (University of Cambridge) and the Creative-Pioneering Researchers Program (Seoul National University). We thank Francis Bloch, Simon Board, Alessandra Casella, Gary Charness, Yeon-Koo Che, Olivier Compte, Vince Crawford, Matt Elliott, Edoardo Gallo, Philippe Jehiel, Navin Kartik, Michael Kosfeld, Boris van Leeuwen, Jay Lu, Rosemarie Nagel, Ryan Oprea, Jacopo Perego, Marzena Rostek, Stephanie Rosenkranz, Tony To, and audiences at a number of seminars for helpful comments.

1 Introduction

Social, economic and infrastructure networks are an important feature of an economy. Over the past two decades, economic theory has explored the role of networks in shaping individual behavior, and the ways in which economic environments shape incentives to create networks.[1] Experimental work on social learning, pricing, trading, public goods, and coordination and cooperation highlights the effects of networks on the efficiency and fairness of outcomes.[2] This work also draws attention to the role of computational complexity and social preferences in shaping individual behavior.[3] With a few exceptions, the bulk of this literature deals with small groups (ranging from four to twelve subjects). In many interesting real world contexts, groups are much larger. As informational demands on individuals and the level of inequality are strongly shaped by scale, to appreciate the scope of these findings, it is imperative that we conduct large scale experiments. This paper presents a new platform that allows us to conduct experiments with up to 100 subjects.

Our platform has a number of novel features. First, it uses a visualisation tool that adjusts the network in real time. This tool relies on force-directed algorithms that aggregate three distinct forces: attraction forces between nodes for visual proximity, repulsion forces between nodes for sparse visualization, and a gravity force attracting every node towards the centre of the screen. The algorithms we adopt in the paper are specially effective in visualising networks that involve very unequal number of connections. Second, we integrate this visualization tool with asynchronous dynamic choice – individuals can form and remove links and change effort levels at any point in time during the experiment. The integration enables us to update rapidly evolving networks in real time on the computer screen. Third, the platform is flexible in information provision both with regard to what subjects know about the network and what they know about the actions and payoffs of different subjects. Finally, the platform allows for both one-sided and two-sided linking, and thus can be used to study a variety of network questions. In this paper, we present three experiments on network formation and assorted activity on this platform.

The paper starts with an experiment titled, 'Linking Game'. In this game, an individual

[1] For a comprehensive overview of the research on networks, see Bramoulle, Galeotti, and Rogers [2016].

[2] For an overview of the experimental literature, see Choi, Gallo, and Kariv [2016] and Breza [2016].

[3] For instance, Chandrasekhar, Larreguy, and Xandri [2019] bring out the role of informational load and cognitive constraints in limiting the extent of information aggregation in social networks. In their work on network formation, Falk and Kosfeld [2012] and Goeree, Riedl, and Ule [2009] argue that inequity aversion inhibits the emergence of unequal networks.

can unilaterally decide to form links with others to access benefits; these links also allow access to benefits that the others have in turn accessed via their links. There is a large literature on such linking games, see e.g., Bala and Goyal [2000], Jackson and Wolinsky [1996], Ferri [2007], Hojman and Szeidl [2008] and Mauleon and Vannetelbosch [2016]. Our experiment is based on the two-way flow model in Bala and Goyal [2000]. A range of network architectures – including the star network and the empty network – can be sustained in equilibrium. The star maximizes aggregate payoffs across the relevant range of parameters (and is therefore efficient).[4] In the star, the hub earns roughly one and a half times the payoffs of the spokes. We conduct an experiment in which the focus is on group size, as it varies from 10 to 50 to 100.

The second experiment is titled, 'Connectors and Influencers'. We enrich the 'Linking Game' with the choice of effort level by players. The value of linking with an individual depends on her level of effort, the links she maintains, and the efforts of those linked to. A number of paper have explored this framework, see e.g., Galeotti and Goyal [2010], Baetz [2015], Perego and Yuksel [2016] and Herskovic and Ramos [2020].[5] The model for our experiment is taken from Galeotti and Goyal [2010]. The equilibrium network is a star; the spokes pay for links with a single hub. There are two effort configurations: a pure influencer outcome (the hub makes all the effort and the spokes choose zero effort), and a pure connector outcome (the hub chooses zero effort, while the spokes choose positive effort). The star structure is efficient, but the level of effort in both effort configurations is too low relative to the first-best (due to the public good aspect of individual efforts). In the pure influencer outcome, the payoffs of the hub and spokes are very similar, while in the pure connector outcome, the hub earns about twice as much as the spokes. In addition to varying the size of groups (respectively, 8, 50 and 100), to examine the role of informational load on subjects, we study a baseline treatment in which subjects are informed only of their own payoffs and a payoff information treatment in which they are informed about everyone's payoffs.

The third experiment is titled, 'Brokerage and Market Power'. We study network formation with trade and intermediation. Linking is two-sided: this distinguishes it from the first two experiments where links could be unilaterally formed. The main treatment variables we consider are pricing rules (criticality and betweennes) and scale (group size

[4]Ferri [2007] shows that star network is the unique stochastically stable network.

[5]These models may be seen as combining the two-way linking model of Bala and Goyal (2000) with the public goods model in networks model of Bramoulle and Kranton [2007].

10, 50 and 100).[6]

The criticality based pricing rule is taken from Goyal and Vega-Redondo [2007]. In this rule surplus from bilateral exchange is divided equally among the two traders and the set of 'critical traders': a trader A is critical for traders B and C in a network if she lies on all paths between the pair. Under this pricing protocol, a wide range of networks – including the hub-spoke network and the cycle network – are pairwise stable.[7] The betweenness based pricing rule is taken from Kleinberg, Suri, Tardos, and Wexler [2008]: surplus in a bilateral exchange is divided equally among the two traders and the intermediaries that lie on the shortest paths between them. This gives rise to intermediation rents in proportion to the betweenness centrality of traders. Under this pricing protocol, a wide range of networks – including the hub-spoke network – are pairwise stable. The cycle, however, is not stable for large groups. As links are costly, the star network is efficient under both pricing protocols. The cycle is almost equally efficient (as it contains only one link more than the star). The star exhibits inequality and this inequality grows with group size; as the cycle is perfectly symmetric, it yields equal payoffs to all subjects. Observe that under criticality pricing, centralized and unequal (such as the star) as well as highly diffused and equal networks (such as the cycle) are stable; this is different from the first two experiments, where theory predicts that (connected) equilibrium networks must be centralized and unequal.

Building on a suggestion of Goeree, Riedl, and Ule [2009], we take the view that, in realistic group sizes, it is more natural to study networks that have the same broad properties as the equilibrium rather than focusing on the creation of an exact equilibrium network. So, for a star, we look for sparseness, inequality of connections, and small average distance, and for a cycle, we look for sparseness, equality of connections, and large average distance. Keeping this perspective in mind, a high level summary of our experimental findings follows.

The first finding is that subjects create sparse networks, in all cases. Moreover, in all but one case, the networks are very unequal and exhibit small average distance.These networks share key properties of the star network and are consistent with the equilibrium

[6]There is a large literature on intermediation: existing work examines pricing by intermediaries, their ability to reduce frictions, and thereby extract surpluses, see Rubinstein and Wolinsky [1987], Condorelli, Galeotti, and Renou [2017], Choi, Galeotti, and Goyal [2017], and Manea [2018]. Condorelli and Galeotti [2016] provide a survey of this work. For experiments on trading in networks and on intermediation, see Gale and Kariv [2009], Kariv, Kotowski, and Leister [2018], Charness, Corominas-Bosch, and Frechette [2007] and Choi, Galeotti, and Goyal [2017].

[7]A network is said to be pairwise stable if no pair of individuals can increase payoffs by adding a link, and if no individual gains by deleting a link (Jackson and Wolinsky [1996]).

predictions. The one exception is the outcome under criticality pricing in the 'Brokerage and Market Power' experiment: here subjects create a network with interconnected cycles – it is equal and exhibits large average distances. This outcome is consistent with the theoretical prediction in that model.

The second finding is that, with one exception, subjects create networks that attain high levels of efficiency – ranging from 70% to 90% of the best equilibrium outcome. The exception is the 'Brokerage and Market Power' experiment with the betweenness pricing rule: in this treatment, for large groups, efficiency level remains below 50%.

The third finding is that scale has powerful effects on networks, efforts and payoffs. In the 'Linking Game' experiment scale has large effects on linking behavior and payoffs: as a result, in small groups, high degree is positively associated with payoffs, while in large groups it is negatively associated with payoffs. In the 'Connectors and Influencers' experiment, scale interacts with payoff information and has large effects on efforts and payoffs. In small groups, there is no perceptible difference in outcomes under the two information treatments: subjects choose the pure influencer outcome. In the baseline treatment, as size grows, the hubs make large investments that result in large losses, while in the payoff information treatment they make low efforts and secure high payoffs. In the 'Brokerage and Market Power' experiment, scale has a major bearing on network structure and earnings. For small groups, there are minor differences in network structure and payoffs under the two pricing rules. By contrast, in large groups, subjects create equal and dispersed networks under criticality pricing and centralized and very unequal networks under betweenness pricing. These findings suggest researchers should be cautious in drawing inferences about behaviour in large scale networks based on evidence from small scale experiments.

Taken together, the experiments illustrate the versatility of our platform: it accommodates experiments involving small as well as a large number of subjects, one-sided as well as two-sided linking, pure linking as well as linking and assorted activity, a variety of pricing protocols, and it allows for variations in information provided to subjects.

The principal contribution of the paper is a new platform to conduct large scale experiments in continuous time.[8] Existing studies on continuous time experiments are built on the development of an experimental software called ConG (Pettit, Friedman, Kephart, and Oprea [2014]) and have focused on small group interaction (see e.g., Friedman and

[8]We hope to make the platform public by Autumn 2020.

Oprea [2012]; Calford and Oprea [2017]). The novelty in our platform is the experimental software that is well suited for studying large-scale network interaction. In order to overcome information overload of evolving networks our software integrates the network visualization tool with the interactive tool of asynchronous choices in real time. This is achieved by adopting an enhanced communication protocol between the server and subjects' computers. It allows us to run both network visualization and asynchronous dynamic choices in real time without communication congestion and lagged responses, even when participants are interacting remotely from different physical locations.

Our experiments are a contribution to the study of networks. They show that moving to larger scale and continuous time has powerful effects that has the potential to significantly reconfigure our perspective on networks.[9] Consider the 'Linking Game': papers by Callander and Plott [2005], Falk and Kosfeld [2012] and Goeree, Riedl, and Ule [2009] present experiments on the same model with small groups (four and six subjects and simultaneous moves). They consider a parametric setting in which the star network is an equilibrium and it is also the efficient network. They find that subjects fail to converge to this network. By contrast, we find that in small groups that are comparable to existing experiments, subjects do indeed form networks that are close to the equilibrium prediction. We attribute this difference to the flexibility afforded by asynchronous choice in continuous time. Our second contribution is to draw out the role of scale and how it interacts with treatment variables like payoff information and pricing rules. As discussed above, these effects of the treatment variables only become evident when we scale up and consider large groups.[10]

We next discuss the relation between our work and the literature on social preferences – specifically the role of efficiency-seeking and inequality-aversion (Charness and Rabin [2002], Fehr and Schmidt [1999], Bolton and Ockenfels [2000], and Kosfeld, Okada, and

[9]There is a related experimental literature on games in networks (Leider, Mobius, Rosenblat, and Do [2009], Charness, Feri, Meléndez-Jiménez, and Sutter [2014], Chandrasekhar, Larreguy, and Xandri [2019]) and on games in which players choose partners and then play a coordination game (Riedl, Rohde, and Strobel [2016], Kearns, Judd, and Vorobeychik [2012]). The interest is on how networks affect behavior and on how allowing for endogenous networks affects behavior. These experiments also involve a relatively small number of subjects (the maximum group size was 36, the only exception is the paper by Leider, Mobius, Rosenblat, and Do [2009]) and they typically also assume simultaneous moves by players.

[10]A similar point can be made regarding experiments on the 'Connectors and Influencers' game. van Leeuwen, Offerman, and Schram [2020] consider a finitely repeated version of Galeotti and Goyal [2010] and report that subjects fail to converge on a pure influencer outcome. By contrast, in our small group experiments, subjects closely match this equilibrium prediction. In addition, our large group experiments present, for the first time, evidence in support of the pure connector outcome.

Riedl [2009]). Falk and Kosfeld [2012] and Goeree, Riedl, and Ule [2009] argue that payoff inequality in the star network is a major factor inhibiting its emergence in the laboratory. How can we reconcile our experimental findings – for instance, in the 'Brokerage and Market Power' experiment the hub earns over thirty times what others earn – with this claim? First, we note that models of inequity aversion are based on average payoff differences and they are relatively 'tolerant' of large inequalities between a few wealthy individuals and the vast majority of population (on this point, also see Schumacher et al. [2017]).[11] This goes some way toward accounting for the large inequality in our experiments. But, building on the work of Charness and Rabin [2002] and Goeree, Riedl, and Ule [2009], we believe that subjects are conscious of efficiency and trade it off against inequality. In our first two experiments, sparse networks with small average distance are efficient and unequal. Subjects unfailingly create unequal networks. In the third experiment, both equal and dispersed networks and unequal and short average distance networks are efficient. Unequal networks (like the star) are always an equilibrium, but equal networks (like the cycle) are also an equilibrium under criticality pricing. Our experiment reveals that when given the choice between equal and unequal equilibrium networks, subjects create equal networks (here 'equal' covers both links as well as payoffs).

Finally, we relate our findings regarding scale effects on individual behavior to the literature.[12] In an influential early contribution, Isaac and Walker [1988] show that there is no pure scale effect in contributions in a public good game. On the other hand, Kagel and Levin [1986] present evidence of more aggressive bidding in auctions with common values, giving rise to a larger winner's curse, as the number of bidders grows. Our findings complement this work: subjects seeking to become a hub form more links and choose higher efforts, as we increase the group size. Interestingly, in the 'Connectors and Influencers' experiment, this happens even when scale does not have payoff implications (in equilibrium, the payoffs to being the hub are invariant with respect to the size of the group). Moreover, scale interacts with information provision in a dramatic manner: in the baseline treatment, increasing scale leads to high efforts and large losses for the hub. By contrast, in the payoff

[11] Recall that in the Fehr and Schmidt [1999] formulation, for a vector of (monetary) earnings $\pi = (\pi_1, \ldots, \pi_n)$, the utility of person i is determined by: $u_i(\pi) = \pi_i - \frac{\alpha_i}{n-1}\sum_j \max(0, \pi_j - \pi_i) - \frac{\beta_i}{n-1}\sum_j \max(0, \pi_i - \pi_j)$ where α defines i's distaste for disadvantageous inequality, and β defines i's distaste for advantageous inequality such that $0 \leq \beta < 1$ and $\beta_i < \alpha_i$.

[12] Outside the economics literature, Centola and Baronchelli [2015] present experimental evidence of scale effects on the emergence of linguistic conventions in naming tasks: in groups with 24 or more subjects, network structure affects convention formation but there is no effect of network structure in groups with 12 subjects.

information treatment, moving from small to large group leads to lower effort and higher earnings for the hub. In the 'Brokerage and Market Power' experiment, scale has powerful effects on network structure, on efficiency and on inequality, as noted above. These findings point to the role of information overload as a first order factor in large scale experiments.

The next section describes our experimental platform. Section 3 presents the experiment on the 'Linking Game', section 4 presents the experiment on 'Connectors and Influencers', and section 5 presents the experiment on 'Brokerage and Market Power'. Section 6 concludes.

2 Experimental Platform

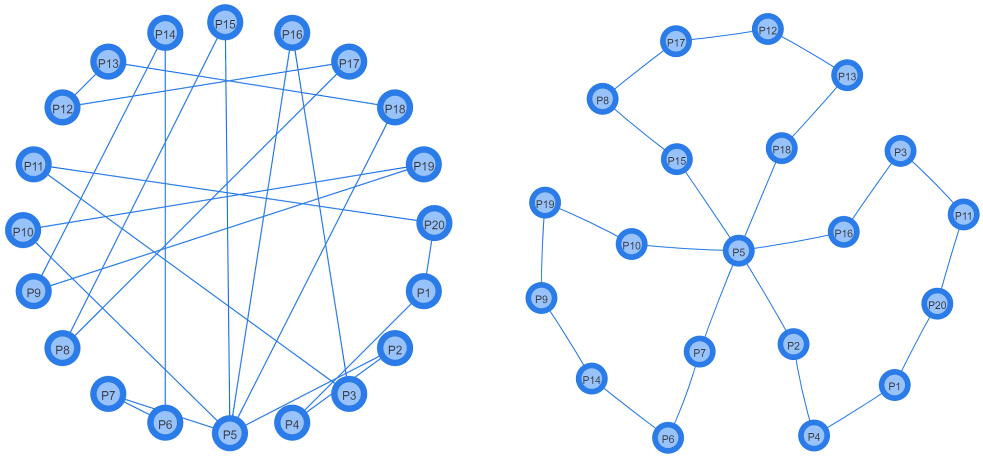
This section discusses four aspects of the experimental platform – network visualization, continuous time asynchronous choices, linking protocols, and information on networks and payoffs.

2.1 Network visualization

Existing studies of network formation in economics have considered small group sizes such as 4 or 8 people and visualized evolving networks with fixed positions of nodes (e.g., Goyal et al. [2017]; van Leeuwen et al. [2020]). When the group size increases, such a representation of networks with fixed positions of nodes makes it very difficult for subjects to perceive network features. For example, consider a group of 20 people with fixed positions of nodes in a circle as depicted in Figure 1a; the exact network is barely perceptible by observing this figure. The same network structure can be represented in a transparent manner in Figure 1b.

For subjects to learn their optimal choices, they must have a good idea of evolving networks. An appropriate tool for visualizing networks and their changes in real time is thus critical in running the experiment in continuous time. This leads us to develop an experimental software including an interactive network visualization tool that allows the network to automatically reshape itself in response to decisions made by subjects. We use *force-directed algorithms* to visualize networks in real time [see, e.g., Eades, 1984, Fruchterman and Reingold, 1991, Hu, 2005, Bostock et al., 2011, Jacomy et al., 2014]. The technical details of the algorithms are provided in Appendix B.

Modelling forces. The force-directed algorithms use attraction and repulsion forces



(a) Fixed visualization (b) Adaptive visualization

Figure 1: Examples of network visualization

between nodes in the network and gravity force toward the center of the screen, in order to readjust their positions in two-dimensional space and improve the overall visibility on the subjects' screen.

Any two nodes o and o' in the network repulse each other with a *repulsion force* $F_r(o, o')$ in order to avoid overlaps and allow a sparse visualization of the network. It is modelled as a decreasing function of the Euclidean distance between two nodes $dist(o, o')$, implying that close nodes repulse more than distant nodes. Two connected nodes o and o' in the network apply an *attractive force* $F_a(o, o')$ towards each other to allow for visual proximity. A classical approach of modelling attraction force is a linear and positive relation with the distance, implying that close nodes attract less than distant nodes. Finally, every node o applies a *gravity force* $F_g(o)$ to the center of the spatialization space O to pull the entire network towards the center of the screen. In particular, such a force allows disconnected components to be within reasonable distance from each other, and therefore more easily visualized on the screen.

The net force vector applied to any node o resulting from the above three forces is then given by the following form of weighted sum (where F_x and F_y represent corresponding force vectors applied to the x and y axes of the Euclidean space respectively):

$$F_x(o) = \frac{x_O - x_o}{dist(o,O)} F_g(o) + \sum_{o' \in N \setminus \{o\}} \frac{x_{o'} - x_o}{dist(o,o')} F_a(o,o') + \sum_{o'' \in N \setminus \{o\}} \frac{x_{o''} - x_o}{dist(o,o'')} F_r(o,o'') \quad (1)$$

$$F_y(o) = \frac{y_O - y_o}{dist(o,O)} F_g(o) + \sum_{o' \in N \setminus \{o\}} \frac{y_{o'} - y_o}{dist(o,o')} F_a(o,o') + \sum_{o'' \in N \setminus \{o\}} \frac{y_{o''} - y_o}{dist(o,o'')} F_r(o,o'') \quad (2)$$

Note that the computation of the repulsion force for every node can be a complex task, especially in the context of large networks. In order to address this issue, the experimental software approximates this computation using the well-known algorithm introduced by Barnes and Hut [1986]. More concretely, it finds groupings of nodes in proximity and determines a repulsion force $F_r(o,c)$ between node o and the group of nodes with a center of mass c, in replacement of the brute force method of computing repulsion forces between all pairs of nodes. More details of this approximation algorithm are provided in Appendix B.1.

We turn back to Figure 1 to derive some intuition of how the net force equations aggregate forces for every node and the network is visualized in the two-dimensional space. The adaptive visualization in Figure 1b is obtained by using the force-directed algorithm. The network has a petal-like structure with three independent sub-components connected through a common player, P5. The visualization algorithm makes P5 to be located at the center of the screen because the neighbors of P5 repluse each other and surround P5, while each pair of P5's neighbors belonging to the same sub-component are in close proximity and positioned side by side. The three forces then operate to make the rest of players located to draw non-overlapping petal-like structures.

Dynamic adjustment. The above equations (14) and (15) describe the net forces that are applied for the visualization of the network, given the positions of all nodes and the links between nodes. When the network changes, the algorithm updates dynamically the network visualization by computing the corresponding velocity of nodes on both coordinate axes.

In order to get a sense of how the network visualization is updated, we turn again to the example of network visualization in Figure 1 and show how the algorithm makes the transition from the fixed visualization in Figure 1a to the adaptive visualization in Figure 1b. Six (slow-motion) snap shots of the transition are presented in Figure 24 in Appendix B. They show how the hub player, P5, moves from the bottom of the fixed circle to the center of the screen, and the petal-like structures emerge. This dynamic adjustment occurs rapidly to arrive at Figure 1b.

In our large-scale experiment, this visualization tool improves graphical clarity of evolving networks and helps subjects distinguish between those who are more connected and those who are less connected. It is wothwhile to note that this tool allows interaction between the subject and the network: while the nodes are subject to the above attraction and repulsion forces, they can also be freely manipulated by the participant through the usual drag-select functionality. The creation and removal of links is also interactive through double-clicking on corresponding nodes. This network visualization tool is built on the open source Javascript library *vis.js*.

2.2 Continuous time with asynchronous choices

It is important to offer subjects adequate opportunities to learn about the environment of decision making, other subjects' behaviors, and how to respond optimally to them. The issues of learning and behavioral convergence can be particularly complicated in a large group. To address them, we build on the work of Berninghaus et al. [2006], Friedman and Oprea [2012] and Goyal et al. [2017], and run the experiment in continuous time with real time updating of all actions and linking by everyone.[13]

Running the continuous time experiments in large groups poses a number of technical challenges. First, every action made by a subject on her computer must be updated instantly on the computer screens of all other participants through the server computer. Network visualization must also be correspondingly updated in real time. As the group size increases, the information flows across the computer network increases dramatically. This can cause communication congestion and lagged responses. Another challenge with a large scale experiment is that it is constrained by the limited capacity of existing laboratories. Large groups that cannot fit into a single lab therefore require remote interactions between subjects in different geographical locations (that is, across different labs). In order to handle both of these technical challenges, we use a Websocket protocol with enhanced

[13] Although the experimental software allows for real time updating of actions, we voluntarily introduce some latency in our experiment to avoid any possible confusion caused by some overload of activity on the subjects' screen. More precisely, the network depicted on any subject's screen is updated every 5 seconds or whenever the subject makes a decision. In a recent paper, Agranov and Elliott [2020] show that experiments with a continuous time protocol lead to a better match with the theoretical predictions as compared to a rigid protocol.

two-way communication between the server and subjects' computers. It fits naturally into the environment of asychronous choices in real time and the updates are made only when necessary. Our Websocket technology relies on the Javascript run-time environment *Node.js.* [14]

2.3 Linking protocols

The experimental platform is flexible enough to accommodate both one-sided and two-sided linking protocols. The network visualization tool presents networks created using one-side linking protocol with directed graphs with arrows on edges. Compared to the one-sided linking protocol, the two-sided linking protocol introduces an extra layer of the relationship between any two individuals: the pair is linked, or unlinked with none of them making a link proposal, or unlinked with only one of them making a link proposal to the other. In order to make it easy for the decision maker to keep track of information on linking relationship in her computer screen, we use a visual representation on the status of the linking relationship between the decision maker, denoted by Me, and an individual as shown in Table 1. An individual who neither made a proposal to nor received a proposal from the decision maker is represented with a circle shape. If an individual sent a link proposal to the decision maker who did not reciprocate it, that individual is depicted with a square shape. If an individual receives a link proposal from the decision maker but did not reciprocate it, the individual is represented with a triangle shape. If both an individual and the decision maker make link proposals to each other, the link between them is visualized with the individual being shaped with a circle.

Figure 2 illustrates how this method of showing different linking relations is added into the network visualization tool. In the initial network depicted on the left side of Figure 2, the decision maker who is represented with a yellow node identified as "me" does not make any link proposal, but receives link proposals from players P2, P3, and P4; these individuals are triangle shaped. From the network on the left, if the decision maker makes link proposals to P2, P3, P4, and P6, the network changes to the right side of Figure 2 (assuming other players do not change their choices). The decision maker then has three realized links with P2, P3, and P4, and one pending link proposal to P6. On the other hand, the decision maker can only see the realized links between any other players (e.g.,

[14] Since it only requires an internet connection and is compatible with most existing web browsers (e.g., Google Chrome, Mozilla Firefox, Internet Explorer), this technology makes no specific restriction on the physical location of every participant.

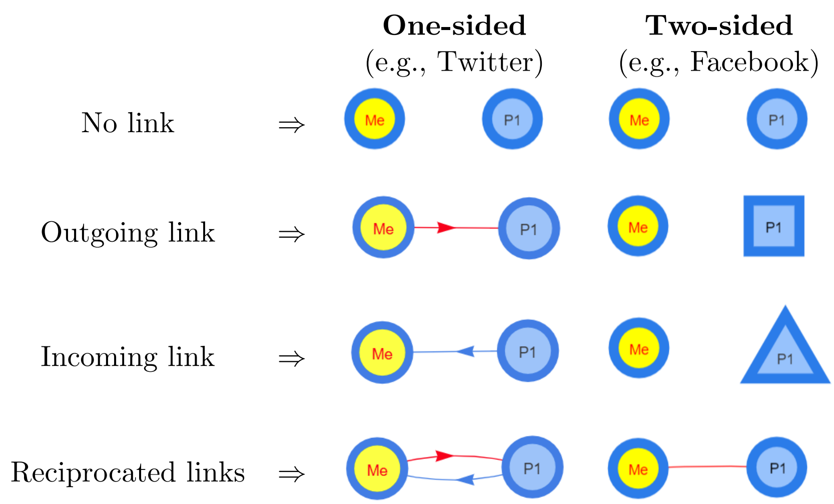


Table 1: Visualization of links

between P1 and P5); no information is provided about unlinked pairs (e.g., the pair of P5 and P7 may be unlinked because either P5, P7, or both P5 and P7 do not make a link proposal).

2.4 Information provision

Our platform is flexible with respect to the level of information that is provided on the network and the payoffs.

Network information. To illustrate the varying degree of network information, consider two extreme scenarios: one, subjects only observe their own neighbors in the current network, and two, subjects get to see the entire network. The information and cognitive load implied by the latter scenario grows rapidly in group size, as illustrated by the network in Figure 3a, which includes 100 players. Thus, there is a potential trade-off between transparency of network visualization and information, and cognitive overload to subjects. Also, different settings of network formation require subjects to observe different amount of network information.

As a means to accommodate for this trade-off, we consider an alternative setting in which subjects are only provided with information about the local structure of the network, within a fixed geodesic distance d. So given a fixed network, for every subject, we can partition the entire group of subjects into two mutually exclusive subgroups: those who

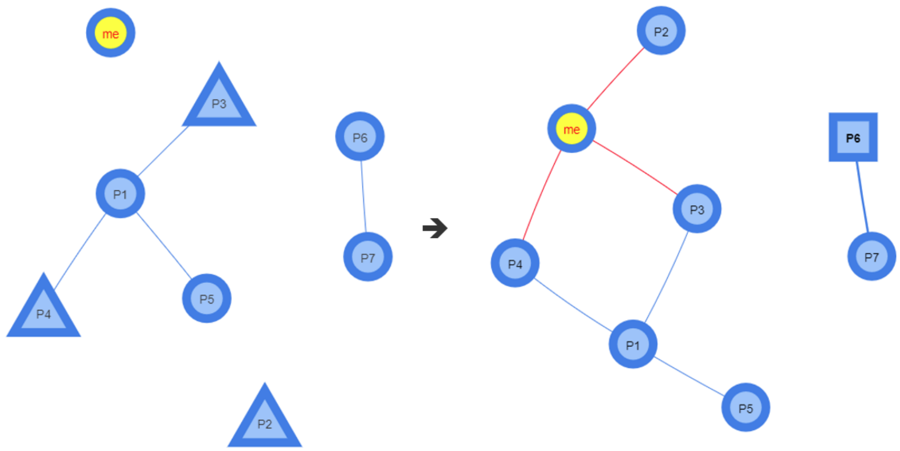
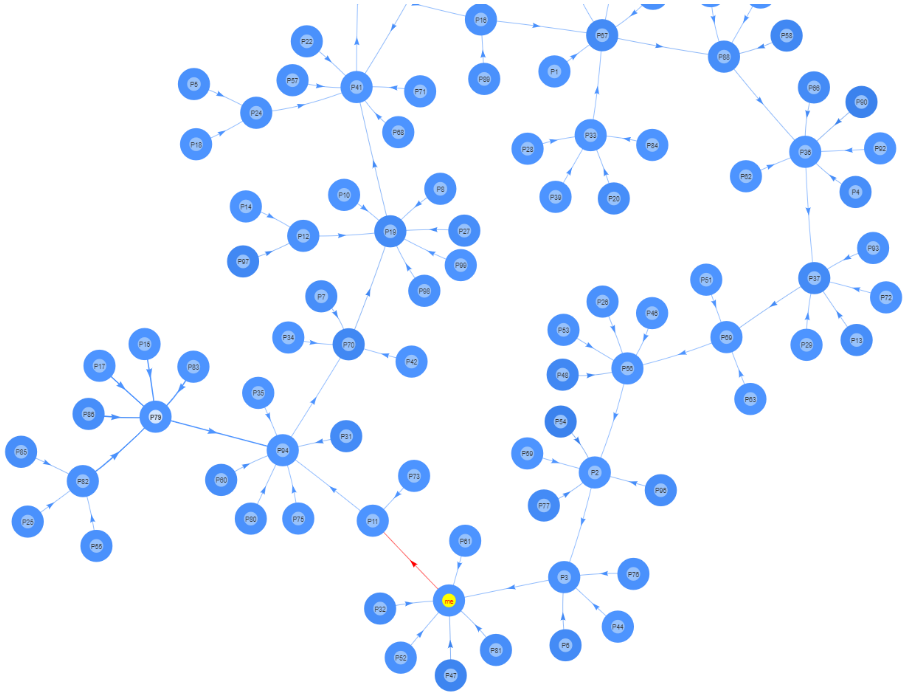


Figure 2: Example of network updates after new choices

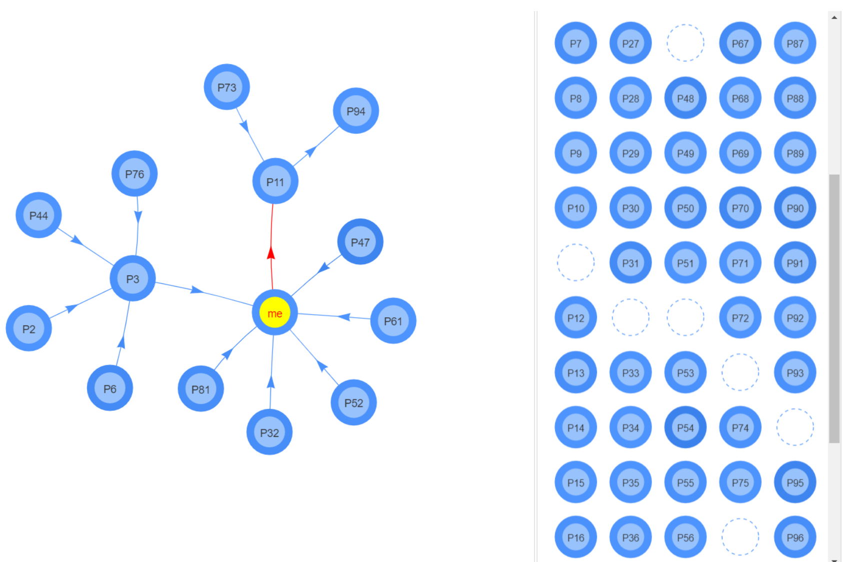
are located within distance d from the subject, and those who are located outside this set. Figure 3b provides an illustration of network visualization and information with 100 subjects. The left side of Figure 3b shows the group of subjects within distance $d = 2$ (and all their links with other subjects within distance 2). The right side of Figure 3b collects the subjects who lie at a distance greater than $d = 2$.

Information on Payoffs. In our experiments, subjects observe their own payoffs in every moment of the game so that they can learn the profitability of their own choices. In principle, the knowledge of others' payoffs could assist subjects in better appreciating the trade-offs associated with different courses of actions, in particular in large groups. This consideration may not be a first order issue in small groups of subjects because subjects can understand payoffs of others in a fairly straightforward manner. However, in a dynamic game with a hundred subjects – and with the network and efforts configuration constantly evolving – an individual may find it much harder to compute the payoffs of other subjects.

The literature of learning in games provides some perspective on this design choice (see Camerer [2003] for a survey). In adaptive models such as reinforcement learning and experience-weighted attraction learning (Camerer and Ho [1999]), players ignore information on payoffs of other individuals. In models of imitation learning (Schlag [1998]) and sophisticated learning (Camerer et al. [2002]), players would behave differently if the payoffs of others are known. When information on others' payoffs is directly available, subjects may follow a different behavioral rule (see e.g., Huck et al. [1999]).

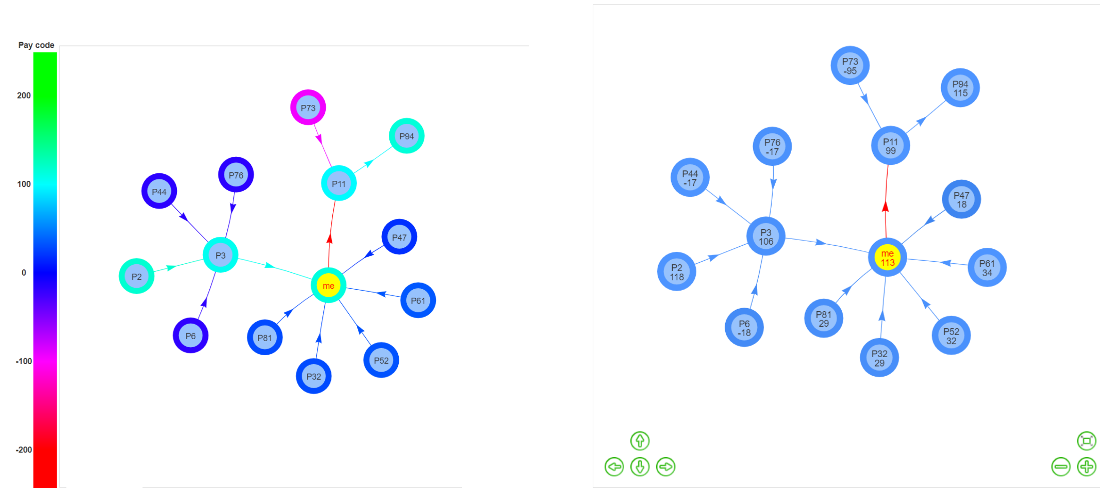


(a) Complete network information



(b) Local network information: d=2

Figure 3: Network information



(a) Color-coded information

(b) Numeric information

Figure 4: Payoff information

Our platform allows for variations on payoff information provision. By way of illustration, we briefly discuss two ways for presenting global payoff information to all subjects. The first option provides such information through a set of color codes, as illustrated by Figure 4a: the color varies from green (high positive payoff) to red (high negative payoff). The scale of the color code is presented on the left hand side at all times during the game. The second option is to present the actual payoffs within the node alongside the player ID, as illustrated in Figure 4b. This may be more effective if the range of payoffs is very large and cannot be accommodated within the colour scheme.

3 The Linking Game

The section presents a simple game of linking: links are unilateral and costly, benefits accrue from individuals accessed through the link and other individuals accessed indirectly via paths in the network. We start with the theoretical model and state a result on (Nash) equilibrium networks. This is followed by a discussion of the experimental parameters and the equilibrium predictions for these parameters. We then discuss the experimental design and procedures. The discussion of findings starts with an overview of the dynamics. We then present the experimental findings on network structure, on efficiency, and on individual behaviour and inequality.

3.1 Theory

We present a model of linking taken from Bala and Goyal [2000]. Let $N = \{1, 2, \ldots, n\}$ with $n \geq 3$. Each player $i \in N$ simultaneously and independently chooses a set of links g_i with others, $g_i = (g_{i1}, \ldots, g_{ii-1}, g_{ii+1}, \ldots, g_{in})$, and $g_{ij} \in \{0, 1\}$ for any $j \in N \backslash \{i\}$. Thus links are unilateral in this game. The set of strategies of player i is $s_i = G_i$, where $G_i = \{0, 1\}^{n-1}$. A strategy profile $s = (s_1, s_2, .., s_n)$ specifies the links made by every player and induces a directed graph, g. Let $\eta_i(g) = |\{j \in N : g_{ij} = 1\}|$ be the number of links i has formed in g. For the purposes of computing benefits, we work with the closure of g: this is an undirected network denoted by $\bar{g}$ where $\bar{g}_{ij} = \max(g_{ij}, g_{ji})$ for every $i, j \in N$. The undirected link between two players reflects exchange of benefits. For any pair of players i and j in g, the geodesic distance, denoted by $d(i, j; \bar{g})$, is the length of the shortest path between i and j in $\bar{g}$. If no such path exists, the distance is set to infinity.

Given a strategy profile s, the payoffs of player i are:

$$\Pi_i(s) = V + \sum_{j \in N} \delta^{d(i,j;\bar{g})} V - \eta_i(g) k \tag{3}$$

where V is the value of benefit per connection, $\delta \in (0, 1]$ is the decay factor associated with indirect access to benefits, k is the cost of linking with another player. We study the Nash equilibrium of this game.

Define a network as efficient if it maximizes the sum of individual payoffs, across the set of all possible networks. The analysis of this model is summarized in the following result.

Proposition 3.1. *A Nash network is either connected or empty. If $k < V(\delta - \delta^2)$ then the complete network is the unique Nash equilibrium. If $V(\delta - \delta^2) < k < V\delta$ the star network is a Nash equilibrium. If $V\delta < k < V(\delta + (n-1)\delta^2)$ the empty network and the star network are both Nash equilibrium. If $k > V(\delta + (n-1)\delta^2)$ then the empty network is the unique Nash equilibrium.*
The unique efficient network is (i) the complete network if $0 < k < 2V[\delta - \delta^2]$, (ii) the star network if $2V[\delta - \delta^2] < k < 2V\delta + V(n-2)\delta^2$ and (iii) the empty network if $2V\delta + V(n-2)\delta^2 < k$.

The game of linking is easy to describe and the equilibrium networks are simple. However, it is worth noting that subjects who wish to choose the best links need to make rather sophisticated computations on the shortest paths to different individuals and the overlap in the neighbourhoods of different players. Over and above this computational difficulty, individuals also confront a very complex coordination problem: in a society with n individuals there are actually n different star networks each corresponding to a different player being the hub; in addition, there is also the empty network. These difficulties grow greatly as we scale up the group size. Thus it is far from clear what networks will actually emerge when individuals choose links and react to the decisions of others on linking.

There are three group sizes in the 'Liking Game' experiment: $N = 10, 50$ and 100. The payoff function is as in equation (3). The value of benefits is $V = 10$ and the decay parameter is $\delta = 0.9$. The costs of linking are adjusted across group size in order to keep the incentives comparable across treatments. The cost of a link is chosen to be $k = 20$ for $N = 10$, $k = 100$ for $N = 50$ and $k = 200$ for $N = 100$.

Given these parameter values, Proposition 3.1 tells us that the empty and star network are both equilibria, for all group sizes. In the star network, the hub and spokes earn 91 and 64, respectively, for $N = 10$, 451 and 308, respectively for $N = 50$, and 901 and 603, respectively for $N = 100$. Thus the star network exhibits significant inequality – the hub earns roughly 50% more than the spokes in all group sizes and the absolute difference in payoffs grows with scale. Individual payoffs in the empty network are equal to 10, in all group sizes. Finally, the star network is efficient for all groups sizes.

Building on a suggestion a suggestion in Goeree, Riedl, and Ule [2009], we focus on general properties of a star network – such as density of links, degree inequality and small average distances. Given that the star is an equilibrium and also efficient, this leads to our first hypothesis.

Hypothesis 1 *Subjects create a network that is sparse, unequal, and has small average distances, for all group sizes, as this reconciles efficiency with individual incentives.*

While the ingredients of the theory are few – the costs of linking and the benefits of linking – and the arguments are simple, it is also clear that, in practice, an individual who is comparing the costs and benefits of forming a link faces a rather complex decision, as she needs to understand what is the return from linking with different individuals and possibly also a combination of many individuals. This requires that she can compute the shortest paths to various individuals in a large and evolving network. Moreover, there exist multiple equilibrium networks – e.g., the empty network and n different star networks (corresponding to different hubs). Indeed, existing experimental studies have concluded that subjects typically fail to conform to equilibrium predictions, see Falk and Kosfeld

[2012] and Goeree, Riedl, and Ule [2009]. So it is far from clear what networks will actually emerge when individuals choose links and react to the decisions of others on linking.

3.2 Experimental design and procedures

The experiment consists of a continuous time game. The game is played over 6 minutes and is referred to as a round. The first minute as a trial period and the subsequent 5 minutes as the game with payment consequences. Every group played 6 rounds.

During a round, at any moment, each subject is informed about the links in their own component and about their own payoff (but not the payoff of any other subject). Figure 25 presents the screen observed by a subject. At any instant in the 6 minutes game, a subject can form/delete a link with any other subject by simply double-clicking on the corresponding node in the computer screen. If the subject forms a link with another subject on the right side of the screen (i.e., someone who is not in the same component), that subject along with the entire component to which they belong would be transferred to the left side of the computer screen. In a case where the subject removes a link with another subject, that subject would be transferred to the right side of the computer screen if they are no longer part of the same component any more and would remain on the left side of the screen otherwise.

At the end of each round, every subject is informed, using the same computer screen, of a time moment randomly chosen for payment. The subject is also provided detailed information on subjects' behavior at the chosen moment, through the corresponding network structure. While the groups were fixed in a session, subjects' identification numbers were randomly reassigned at the beginning of every round in order to reduce potential repeated game effects. The first round was a trial round with no payoff relevance and the only the last 5 rounds were relevant for subjects' earnings. In analyzing the data, we will focus on subjects' behavior and group outcomes from these last 5 rounds. This procedure of the experiment and data usage is common to the three experiments.

A subject participates in only one of the experimental sessions.[15] After subjects read the instructions, the instructions were read aloud by an experimenter to guarantee that they all received the same information. While reading the instructions, the subjects were provided with a step by step interactive tutorial which allowed them to get familiarized with the experimental software and the game. Subjects interacted through computer terminals and

[15]Subjects also participated in only one of the experiments presented here.

the experimental software was programmed using HTML, PHP, Javascript, and SQL. This procedure is common to all three experiments reported in the paper. Sample instructions and interactive tutorials for all the experiments in the paper are available in Appendix C.

In total there were 9 sessions: 1 session of 4 groups of 10 subjects for the $N = 10$ treatment, 4 sessions of 50 subjects for the $N = 50$ treatment, and 4 sessions of 100 subjects for the $N = 100$ treatment. In each experimental session, subjects were matched to form a group and interacted with the same subjects throughout the experiment. This protocol of group formation is also common to all three experiments. Therefore, there are 4 independent groups for each group size. A total of 640 subjects participated in the experiment.

At the beginning of the experiment, each subject was endowed with an initial balance of 50 points in the $N = 10$ treatment, 250 points in the $N = 50$ treatment, and 500 points in the $N = 100$ treatment. Subjects' total earnings in the experiment were equal to the sum of earnings across the last 5 rounds and the initial endowment.[16] Earnings were calculated in terms of experimental points and then exchanged into euros at the rates of 40 points being equal to 1 euro for the $N = 10$ treatment, 200 points being equal to 1 euro for the $N = 50$ treatment, and 400 points being equal to 1 euro for the $N = 100$ treatment.[17] On average, a session lasted 90 minutes. On average subjects earned 15.3 euros (this includes a 5 euros show-up fee).

All three experiments reported in the paper were conducted in the Laboratory for Research in Experimental and Behavioral Economics (LINEEX) at the University of Valencia. The experimental sessions of the $N = 100$ treatment were conducted through the internet connection between LINEEX and the Laboratory for Experimental Economics (LEE) at the University Jaume I of Castellón. In this case the number of subjects was then evenly distributed across the two locations. Subjects in the experiment were recruited from online recruitment systems of the LINEEX and LEE.[18]

[16] In case of negative total earnings, the corresponding subject would simply earn 0 point from the game.

[17] The different conversion rates and initial endowments are justified by the different linking costs across different treatments, as an attempt to maintain similar earnings.

[18] At the end of each experiment, subjects took incentivized tasks to elicit social preferences and risk preferences. They are a modified version of Andreoni and Miller [2002] and Holt and Laury [2002], respectively. In addition, subjects answered a brief version of the Big Five personality inventory test adapted from Rammstedt and John [2007], a comprehension test related to the corresponding experimental game, and a debriefing questionnaire including demographic information. More details about them can be found in Appendix E.

3.3 Results

3.3.1 Overview

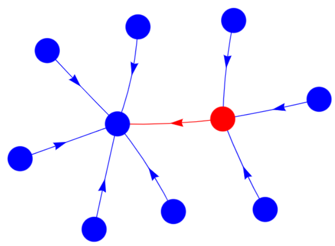
We begin by presenting snap shots taken from the experiments with ten subjects and a hundred subjects. While coming from a particular group, these snapshots are representative of the dynamics in the experiment. Starting with a group of $n = 100$, Figure 5c shows that one subject (shown in red) emerges as a hub at minute 3. The hub status of this subject is maintained by links formed by other subjects as well as those formed by herself (also shown in red). At the end of the game (Figure 5d), the same subject remains as a hub. However, by the end, this subject has now deleted all the links she created and maintains the hub status only by attracting more links from other subjects. We observe similar dynamics in a small group of $n = 10$ in Figures 5a and 5b.

The key points from these snap shots are: first, there is specialization in linking so that the emerging network is sparse and unequally connected. Second, there is a big difference in the linking investments of the hub in the small group and the large group: in the small group the hub forms few links, while in the large group, the hub forms a very large number of links to attract links from others. This suggests that, in large groups, the hub invests at the early stages and hopes to recover these costs by becoming the hub (and eventually deleting all links). The data analysis examines these points systematically.

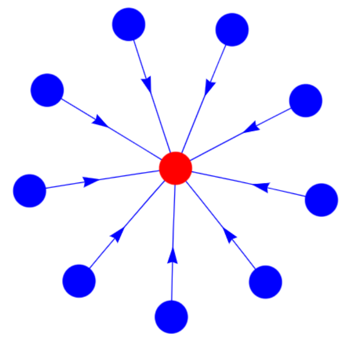
For simplicity, in the data analyses that follow, the data used from every round of the game consists of 360 observations (snapshots of every subject's choices) selected at intervals of one second. Although some information about choice dynamics between two time intervals may be lost, we believe that the impact of such a simplification is so small as to be negligible. Moreover, unless stated otherwise, all statistical analyses consider data from the last 5 minutes (for each round of the experiment).

3.3.2 Network structure

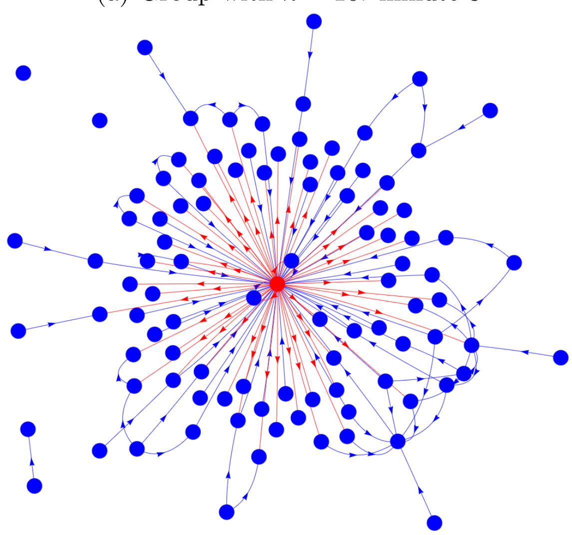
The first set of data analyses of interest are the dynamics of network structure observed during the game, across different group sizes. We focus on four properties of the network structure: (*i*) network sparsity, (*ii*) inequality of linking, (*iii*) network closeness, and (*iv*) the stability of network structure. We use average per capita indegree as a measure of sparseness of a network, the Gini coefficient of indegrees as a measure of inequality of linking, average distance between two nodes and closeness centrality as measures of network



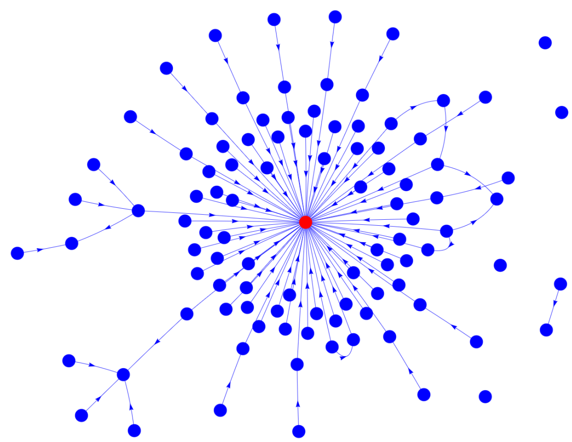
(a) Group with $n = 10$: minute 3



(b) Group with $n = 10$: minute 6



(c) Group with $n = 100$: minute 3



(d) Group with $n = 100$: minute 6

Figure 5: Snap shots of Linking Game: $n = 10$ and $n = 100$

closeness, and the persistence of a hub over time and the per capita number of decisions made as measures of the stability of a network.

Figure 6 summarizes our analysis of the network structure.[19]

Firstly, subjects across all group sizes create sparse networks. Figure 6a shows that in the $N = 10$ group, the average indegree is around 1, and that it is stable from the beginning until the end of the game. In the large groups of $N = 50$ and $N = 100$, at the start, the average indegree goes over 2, but gradually comes down and is close to 1 by the end of the

[19]The vertical dotted line in figures of time series represents the beginning of the payoff-relevant part of the game, i.e. the last 5 minutes.

game. Recall that average indegree of 1 is the minimum needed to ensure connectivity of a network. Thus, we conclude that subjects create sparse networks in all group sizes.

Secondly, subjects create very unequal networks. Figure 6b plots the Lorenz curves of indegree across different group sizes; recall that the Lorenz curve reflects the cumulative fraction of subjects, ranked from least connected to most connected, against the cumulative fraction of total indegrees.[20] Figure 6b also shows the corresponding Gini coefficients of indegree across group sizes: 72% for $N = 10$, 74% for $N = 50$, and 69% for $N = 100$.[21] They reveal that speacialization is strong and present in every group size.

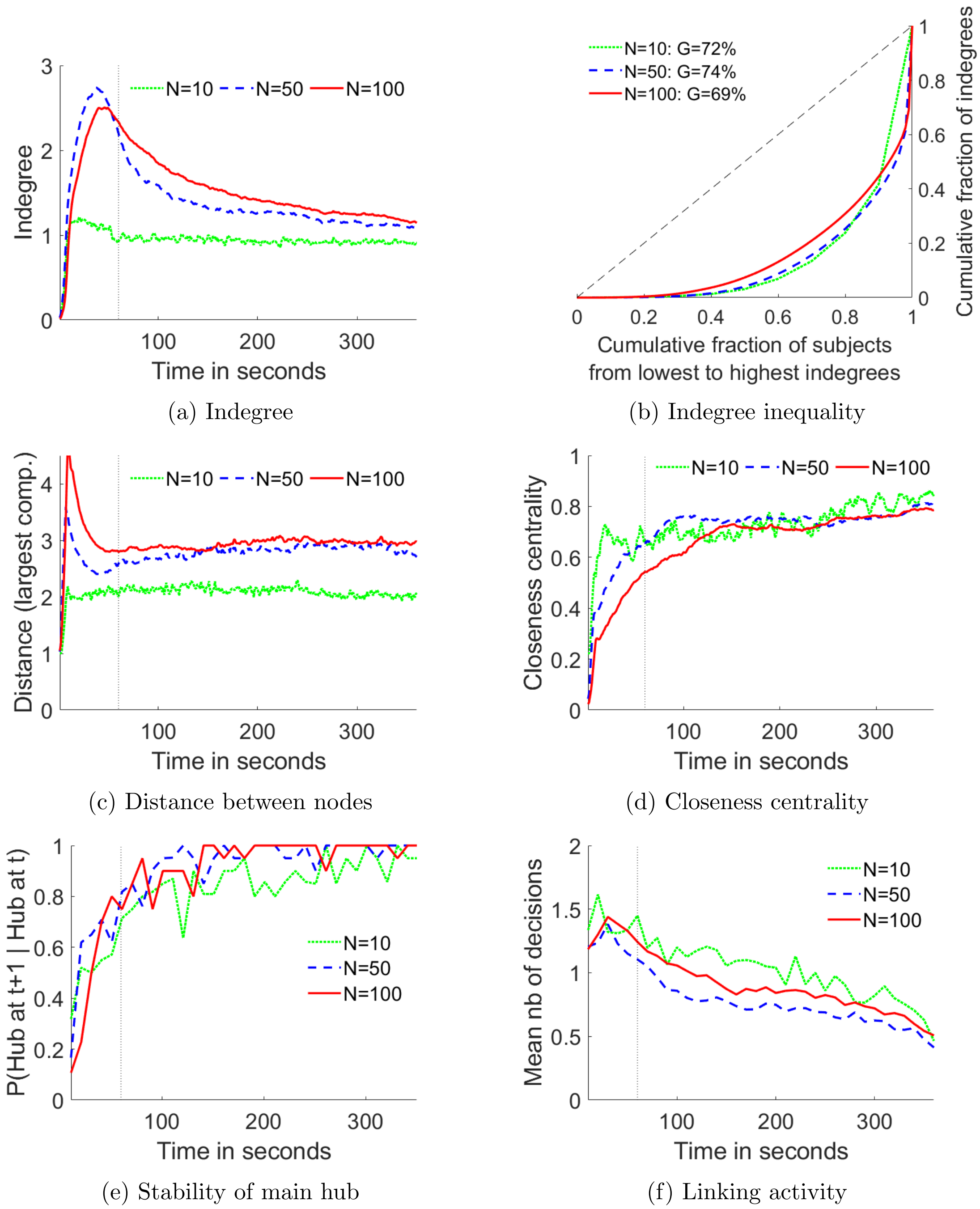
Thirdly, subjects create networks that have small average distances. Figure 6c present the dynamics of average distance in the largest component of networks.[22] The average distance in the $N = 10$ treatment is 2 and stable, whereas those in the large groups are close to 3. Thus, although there is a group size effect, the average distance in all group sizes is rather small. In addition, following a suggestion of Goeree, Riedl, and Ule [2009] we look at a general measure of centrality – closeness centrality – and its dynamics over time across group sizes. Closeness centrality compares the proximity between players in a network with that of the star (see Appendix F for the formal definition of closeness centrality used in Figure 6d). In all group sizes, closeness centrality is between 0.7 and 0.8: in our view this is very close to the star network. In comparison, simulations of well known networks generate significantly lower closeness centrality: 0.52 by generating 1000 scale free networks with $N = 100$ based on Barabási and Albert [1999]; 0.22 by generating 1000 random networks with $N = 100$ and an average degree of 3 based on Erdos and Rényi [1959]. We conclude therefore that subjects are successful at creating networks with very small average distance and high closeness centrality, across all groups sizes.

Finally, we study stability and convergence of networks. The first measure pertains to the persistence of hub status: we consider the time interval of ten seconds and define as the main hub the person who is most connected for the longest duration over that interval. We then check whether the main hub maintains its status in the next 10-second interval. Figure 6e plots the time trends on the probability of persistence. In all three groups the

[20] The Lorenz curve and the corresponding Gini coefficient are first computed for every second, and averaged across the last 5 minutes, in all rounds and groups.

[21] There is no statistical difference of the Gini coefficients between any two group sizes by using the group-level average data (two-sample t-test: p=0.51 for $N = 10$ and $N = 50$; p=0.53 for $N = 10$ and $N = 100$; p=0.16 for $N = 50$ and $N = 100$).

[22] The average size of the largest component is close to the group size in each treatment: 9.3 in the $N = 10$ treatment, 48.4 in the $N = 50$ treatment, and 96.3 in the $N = 100$ treatment.



(a) Indegree

(b) Indegree inequality

(c) Distance between nodes

(d) Closeness centrality

(e) Stability of main hub

(f) Linking activity

Figure 6: Network Structure

stability of the main hub increases rapidly in the first two minutes and eventually is over 80%. Thus, once a hub emerges, its status is very stable. Moreover, in the large groups, this persistence is especially strong: it rises fast and is close to 100%. In addition to the study of the stability of the hub, we study if the overall network is converging. Following Goeree, Riedl, and Ule [2009], we plot the per capita number of link changes over time. Figure 6f shows a clear decline in this measure: the average number of decisions made by a subject per block of 10 seconds is close to 0.5 toward the end of the game in all groups.

Result 1.1 In the Linking Game experiment, subjects create sparse, unequal, and small average distance networks in all group sizes. These network properties are broadly consistent with Hypothesis 1.

3.3.3 Efficiency

We study efficiency by considering, at each second, the ratio of observed total payoffs to the best Nash equilibrium payoff, i.e. the payoff in the star network. Figure 7 plots the dynamics of this ratio across different group sizes. In the $N = 10$ group, we observe that subjects attain 80% of efficiency after the first one minute and the level of efficiency steadily increases to reach around 90%. In the large groups the level of efficiency is around 40% after the first one minute, but increases rapidly in the next one minute and then steadily through the rest of the game. At the end of the game, this ratio is close to 80% in the $N = 50$ treatment and to 75% in the $N = 100$ treatment. The differences in efficiency across group sizes are statistically significant: two-sample t-test with the group level average data: $p < 0.01$ for $N = 10$ and $N = 50$, $p < 0.01$ for $N = 10$ and $N = 100$, and $p < 0.05$ for $N = 50$ and $N = 100$. We summarize these findings as follows.

Result 1.2 In the Linking Game experiment, subjects create networks that attain high levels of efficiency, across all group sizes. The level of efficiency attained is falling in group size.

3.3.4 Behavior and inequality

This section examines the effects of group size on individual behavior and payoffs. The snap shots in Figure 5 suggest that there are different types of subjects – a few subjects who form many links in a bid to become the hub, while the rest of the subjects simply link

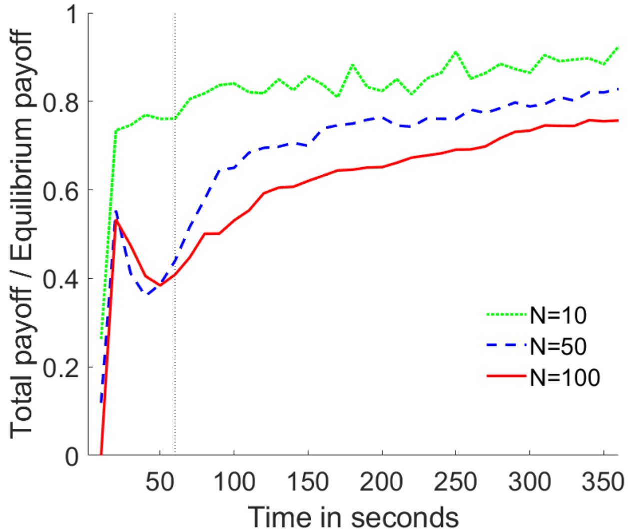


Figure 7: Efficiency in the Linking Game

with the most connected individual. This suggests, at every second, a simple classification of subjects into three categories – most connected, 2nd most connected, and the others.

Figure 8 plots the time series of the number of links created – the out-degree – by each type, normalized by the total number of feasible links $(n-1)$. It also plots the time series of payoffs received by each type of subject.

Group size has powerful effects on linking behavior for the two most connected individuals. In the small group, the two most connected subjects create very few links – indeed they create fewer links than the other subjects. In contrast, the most connected subject in the large groups forms a very large number of links at the early part of the game, but deletes most of them by the end of the game. We observe a similar pattern of linking for the second most connected individual, albeit to a lesser extent. The rest of subjects in the large groups create very few links, over time.[23]

These dynamics of linking have strong implications for individual payoffs. In the large groups, in the early part of the game, the two most connected subjects make negative earnings because of over-linking activities. The most connected subject receives the largest earnings toward the end of the game, but it is clear that average earnings of the hub across

[23]Differences in behavior and payoffs across different types of subjects are supported by corresponding regression analyses provided in Appendix G.1. In all regression analyses reported throughout the paper, we include fixed effects for rounds and groups.

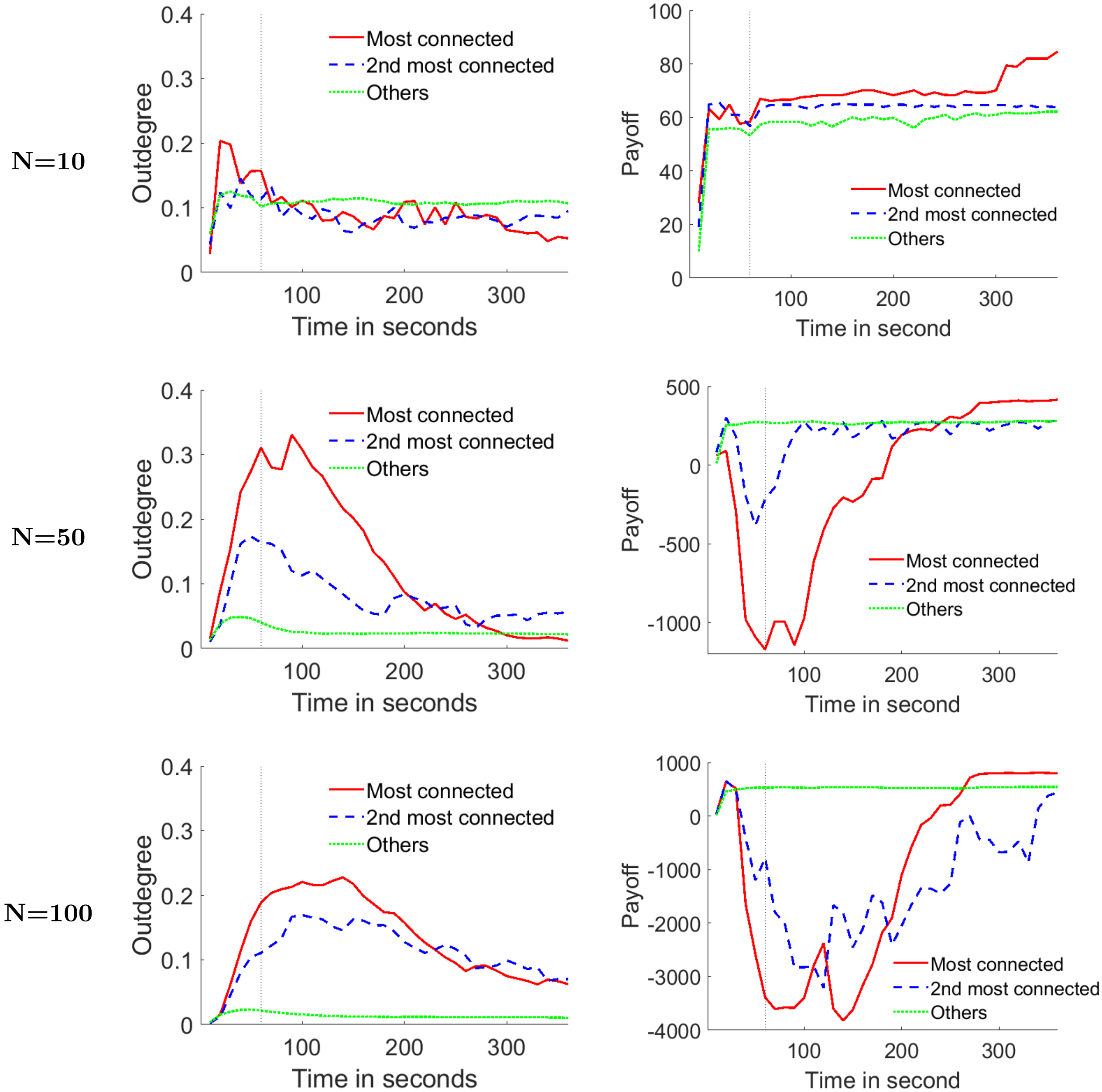


Figure 8: Behavior and inequality.

time are very low. The competition to become hubs is much less pronounced in small groups; as a result, the most connected individuals perform better than the other subjects on average, across time.

We summarize these findings on individual behavior and payoffs as follows.

Result 1.3 In the Linking Game experiment, group size has powerful effects on individual linking and payoffs. In small groups, the hub forms few links and earns more than spokes. In large groups, a few subjects make a very large number of links in a bid to become the hub; as a result, they earn less than the other subjects. Higher degree is associated with higher payoffs in the small group and with lower payoffs in the large groups.

The Linking Game experiment illustrates the use of our platform in a particularly simple setting. It is worth noting a few aspects of the experimental findings. Subjects create sparse networks, with unequal connections and small average distance in all cases. But the dynamics of behaviour are very different as we increase group size: in particular as the last result above indicates, the behaviour of the hub and their competitors changes dramatically. Their behaviour poses a puzzle, as this intense linking activity leads to lower average payoffs as compared to the rest of the subjects. We will take up this issue after we present the findings of the next experiment.

4 Connectors and Influencers

This section presents an experiment on a game in which individual form links and also choose effort levels. The second experiment therefore builds in two ways on the first experiment: by expanding the strategic possibilities and by varying the information level available to subjects.

4.1 Theory

We enrich the pure linking model to incorporate efforts by individuals. This model of linking and efforts is taken from Galeotti and Goyal [2010]. Each player $i \in N$ now simultaneously and independently chooses a level of effort $x_i \in \mathbf{R}_+$ and a set of links g_i with others to access their efforts such that $g_i = (g_{i1}, \ldots, g_{ii-1}, g_{ii+1}, \ldots, g_{in})$, and $g_{ij} \in \{0, 1\}$ for any $j \in N \backslash \{i\}$. Let $g_i = \{0, 1\}^{n-1}$. We define the set of strategies of player i as $s_i = \mathbf{R}_+ \times G_i$, and the set of strategies for all players as $S = S_1 \times \ldots \times S_n$. A strategy profile $s = (x, g)$ specifies efforts and the links made by every player. Define $N_i^l(\bar{g}) = \{j \in N : d(i, j; \bar{g}) = l\}$ to be a set of players at distance l from i in $\bar{g}$.

Given a strategy profile $s = (x, g)$, the payoffs of player i are:

$$\Pi_i(x, g) = f(x_i + \sum_{l=1}^{n-1} a_l \sum_{j \in N_i^l(\bar{g})} x_j) - cx_i - \eta_i(g)k \tag{4}$$

where c denotes the constant marginal cost of efforts, k the cost of linking with another player, and a_l reflects the spillover across players who are at distance l. So if $j \in N_i^l(\bar{g})$, then the value of agent j's effort to i is given by $a_l x_j$. Throughout, it is assumed that $a_1 = 1$, $a_2 \in (0, 1)$, and $a_l = 0$, for all $l \geq 3$. The benefit function $f(y)$ is twice continuously differentiable, increasing, and strictly concave in y. For simplicity, also assume that $f(0) = 0$, $f'(0) > c$, and $\lim_{y \to \infty} f'(y) = m < c$. Under these assumptions, there exists a number $\hat{y} \in X$ such that $f'(\hat{y}) = c$.

The analysis of Galeotti and Goyal [2010] focuses on polar cases in which $a_1 = 1$ and $a_l = 0$, for all $l \geq 2$ and the case where $a_l = 1$, for all l. Our formulation allows for indirect flow of benefits with decay. We provide a characterization of Nash equilibrium for this case, when the linking costs are relatively large.

Proposition 4.1. *Suppose payoffs are given by (4), and that $a_1 = 1$, and $a_2 \in (0, 1)$. Then there exists a $\hat{k}$, such that for $k \in (\hat{k}, c\hat{y})$ the following is true. The equilibrium network is a periphery sponsored star. There exist two possible effort equilibrium configurations:*

- *the pure influencer outcome: the hub invests $\hat{y}$ and everyone else invests 0.*
- *the pure connector outcome: the hub invests 0 and everyone else invests $\hat{y}/(1 + (n - 2)a_2)$.*

The proof is provided in Appendix A. In the pure influencer equilibrium, we see an extreme version of the 'law of the few': a single person receives all the links formed in society and also carries out all the efforts. The pure connector equilibrium retains the specialization in links: a single person receives all links, but the efforts are evenly spread out. Interestingly, in both equilibria the creation of links is basically egalitarian: $n - 1$ players each form one link. For large k values, the payoff distribution is only slightly unequal in the pure influencer equilibrium. However, the payoff inequality can be very large in the pure connector equilibrium between the hub and every spoke (especially if k

is large and a_2 is small). The pure connector equilibrium holds only for a sufficiently large group size n, i.e., $n \geq 2 + k/(a_2(c\hat{y} - k))$.

We now specify the parameters used in the experiment. The function $f(.)$ is taken from Goyal et al. [2017] and as follows:

$$f(y) = \begin{cases} y(29 - y) & \text{if } y \leq 14 \\ 196 + y & \text{else} \end{cases} \tag{5}$$

For simplicity, the efforts are assumed to take on integer values only and there is an upper bound, $\overline{x} = 20$. The set of efforts is given by $X = [0, 20]$. The cost of effort $c = 11$ and the cost of a link $k = 95$; finally, the decay parameter $a_2 = 1/2$. Given these parameters, the stand alone optimum effort, $\hat{y} = 9$.

We will consider groups of size 8, 50 and 100. Given these parameter values, Proposition 4.1 tells us that, for all groups sizes, there exists a pure influencer equilibrium in which a single individual chooses 9, all other individuals choose 0 and form a link with the positive effort player. Moreover, given the integer constraints, the minimum positive effort is 1. There is no pure connector equilibrium for $n = 8$. However, in the treatments with 50 and 100 subjects, the star where 18 peripheral individuals choose 1 and the rest of the subjects choose 0 constitutes an 'approximate' pure connector equilibrium (for details see Appendix A).[24]

In the pure influencer equilibrium, the hub chooses effort 9, while the spokes choose 0. The hub earns 81, while the spokes each earn 85. In the pure connector equilibrium, the hub chooses effort 0, eighteen spokes choose 1 each, while the other spokes choose 0. The hub earns 198, the active spokes 74, and the inactive spokes 85. Hence, there is little inequality in the pure influencer equilibrium, but significant earnings inequality in the pure connector equilibrium.

As the costs of effort are linear and there is distance based decay, for any given level of effort, the hub-spoke network maximizes aggregate player welfare. Thus the star is the efficient network architecture.

Taking together our characterization of equilibrium with our observations on efficiency and equity leads us to the following hypotheses:

Hyopothesis 2A *Subjects create networks that are sparse, unequal, and have small dis-*

[24] The periphery player who chooses effort 1 and forms a link with the hub earns 79.25. This person could earn 81 by deleting the link and instead choosing effort level 9.

tances, for all group sizes, as that reconciles efficiency with individual incentives.

Hyopothesis 2B *Subjects choose the pure influencer configuration of efforts, as that reconciles efficiency with equity and individual incentives, for all group sizes.*

We conclude the theory by noting that individuals need to make a difficult set of computations, as they need to keep in mind the effort levels and the entire network of connections to make comparisons of returns to linking across different potential nodes. As anyone can serve as a hub, there are as many star networks as number of players. Over and above this, in the large groups, in addition to the pure influencer equilibrium, there is also the pure connector equilibrium, so the coordination problem faced by individuals is indeed formidable. Indeed, existing experimental studies on this game have concluded that subjects typically fail to conform to equilibrium predictions, see e.g., van Leeuwen, Offerman, and Schram [2020].

4.2 Experimental design and procedures

The experiment varies the size of groups and the information on others' payoffs. There are three group sizes – 8, 50 and 100. Every subject observes the network within distance $d = 3$. This is motivated by the payoff structure in this game: recall, that every player can access benefits from their neighbours and the neighbours of their neighbours. So, in the experiment, in order to understand the incentives for effort of a neighbour a subject needs to be able to see the relevant neighbourhood of this neighbour. In the baseline treatment, subjects observe only their own payoffs: they see the screen as presented in Figure 26. In the payoff information treatments, called PayInfo, subjects not only observe their own payoffs but also the payoffs of all other players. In the PayInfo treatment subjects are shown the payoffs of others through variations in the colours of the nodes (as illustrated in Figure 4a). Therefore, the experiment consists of 6 treatments: 3 group sizes × 2 information treatments.[25]

The number and duration of rounds is exactly as in the 'Linking Game'. There are two substantive differences in the design: the first one pertains to network information one – subjects only see the local network, up to distance 3 neighbourhood. This leads us to use a slightly different design. If the subject forms a link with another subject on the right side

[25]We also conducted an experiment with 4 subjects: the outcomes were very similar to the outcomes with 8 subjects. These results are not reported to save space but available upon request from the authors.

of the screen (i.e., someone who is in more than 3 geodesic distance away), that subject along with her neighbors and neighbors' neighbors would be transferred to the left side of the computer screen. In a case where the subject removes a link with another subject on the left side of the screen, that subject would be transferred to the right side of the computer screen if they go more than 3 links apart and would remain in the left side of the screen otherwise. The second difference is that subjects also choose any level of effort: they do this my moving a slider varying from 0 to 20 by increments of 1. This slider is provided on top of the decision screen along with other payoff-relevant information including the subject's gross earnings (i.e., the benefit $f(x)$ where x is the total amount of information the subject has access to), cost of effort, cost of linking, and resulting earnings (i.e., payoff $\Pi_i(x_g)$). See Figure 26 in Appendix D for an illustration.

In total there were 12 sessions: 1 session with 4 groups of 8 subjects for each of the Baseline8 and PayInfo8 treatments, 4 sessions of 50 subjects for each of the Baseline50 and PayInfo50 treatments, and 3 sessions of 100 subjects for each of the Baseline100 and PayInfo100 treatments. A total of 1064 subjects participated in the experiment.

At the beginning of the experiment, each subject was endowed with an initial balance of 500 points and added positive earnings to or subtracted negative earnings from that initial balance. Subjects' total earnings in the experiment were the sum of earnings across the last 5 rounds and the initial endowment. Earnings were calculated in terms of experimental points and then exchanged into euros at the rate of 100 points being equal to 1 euro. Each session lasted on average 90 minutes, and subjects earned on average about 18 euros (including a 5 euros show-up fee).

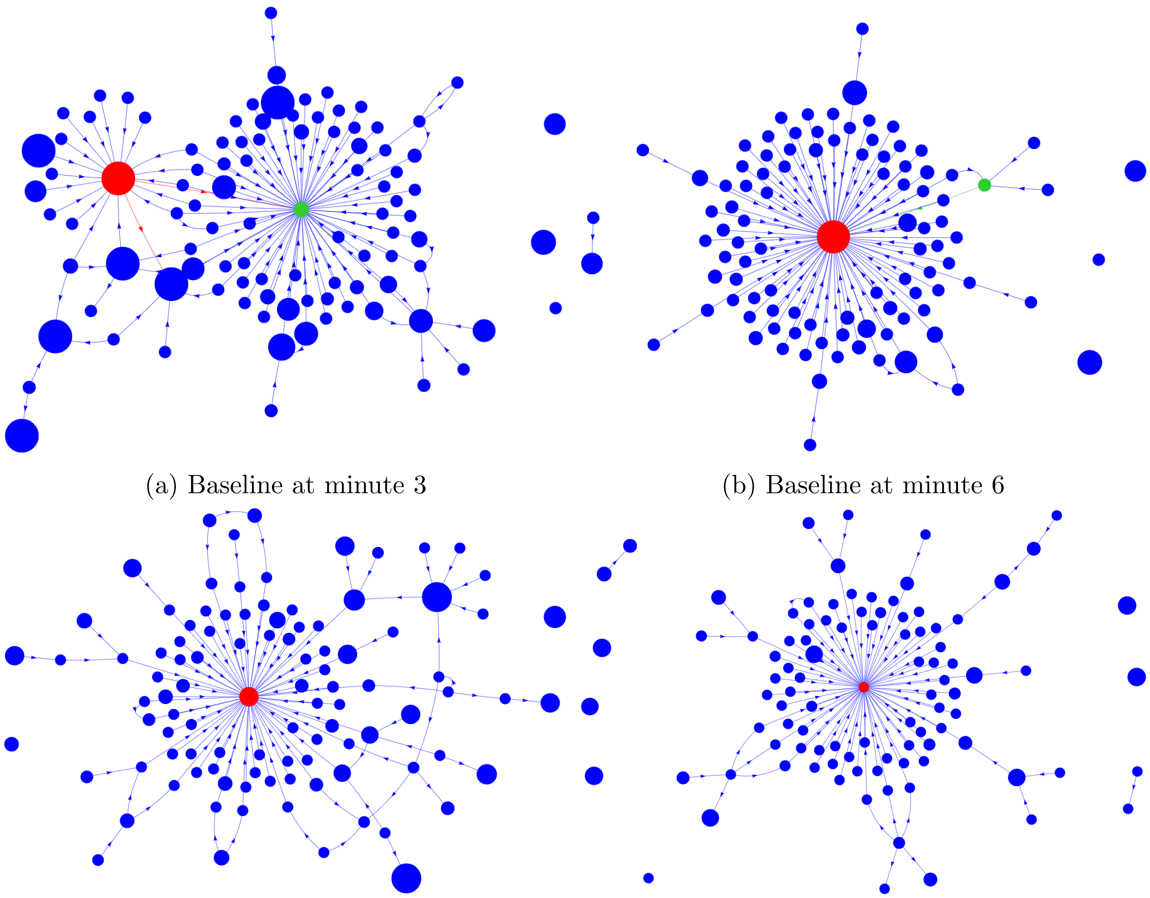
4.3 Results

4.3.1 Overview

We start with the snap shots taken from the experiment with a hundred subjects in the baseline and in the payoff information treatment. Figure 9 presents networks emerging at minute 3 and at minute 6 in both the baseline and the payoff information treatment. In this figure, the node size characterizes the individual effort level of the corresponding player subject (the larger the node, the higher the individual effort). The nodes colored in red and green identify the same subjects across the two different moments (e.g., the red node in Figures 9a and 9b corresponds to the same subject in the group).

In the baseline treatment, we observe the initial emergence of a main hub at minute

3 (green player) who, after attracting many connections, lowers her effort. This hub is however being challenged by a secondary hub (red player) who makes maximal effort. As a result of this competition, the green player gradually loses her links to the red player over time, until the red player is the only hub in the network, as shown at minute 6.



(a) Baseline at minute 3

(b) Baseline at minute 6

(c) PayInfo at minute 3

(d) PayInfo at minute 6

Figure 9: Snap shots from Baseline100 and PayInfo100 treatments (node size characterizes the individual effort level of the corresponding player)

Figures 9a and 9b draw attention to three points: one, extreme specialization in linking and efforts; two, intense competition among a few subjects to become the hub, reflected in very large efforts; and three, the emergence of the pure influencer outcome.

Figures 9c and 9d present snapshots taken in the payoff information treatment. Specialization in linking continues to hold in this setting. However, there is a major change in the behavior of individuals seeking to become a hub: the most connected individual (red player) starts at a high effort, but then shades her efforts. The key difference with the baseline is that no other player challenges her effectively for the hub position. The outcome is closer to the pure connector outcome in this case, where the most connected individual chooses zero effort and forms no link, and consequently earns much more than the peripheral individuals.

We now turn to a systematic analysis of the experimental data.

4.3.2 Network structure

Figures 10 and 11 summarize the properties of network structures in each of the information treatments.

Figures 10a and 11a show that subjects create sparse graphs: average indegree is less than 1 in the small groups in both baseline and payoff information treatments. In the $N = 50$ group, for both information treatments, the average indegree is stable around 1 over time. In the Baseline100 treatment, average indegree is falling over time to reach 1 at the end of the game. The average indegree is steadily increasing in the PayInfo100 treatment, but remains below 1.5. Recall that in the star network, the average indegree would be slightly below 1. Thus subjects create sparse networks.

Figures 10b and 11b present the Lorenz curves and the corresponding Gini coefficients of indegrees across different group sizes. They reveal that link distribution is unequal in every group size, but that it becomes especially acute in the baseline treatment as the group size increases. The Gini coefficient is 70% for Baseline8, 86% for Baseline50, and 89% for Baseline100. By organizing the group-level average data, we observe a statistical difference in this measure of inequality in linking between the small groups ($N = 8$) and any of the large groups ($N = 50$ or $N = 100$) at the 5% significance level. In the payoff information treatments, the Gini coefficients are 77% for PayInfo8, 84% for PayInfo50, and 81% for PayInfo100. We conclude that subjects create very unequal networks, and this inequality grows with group size in the baseline treatment.

Figures 10c and 11c show that average distances in the small group of both information treatments is no larger than 2. The average distance in the large groups under both information conditions converges to around 3. Recall that the average distance in a star network would be close to 2. We conclude that the average distance is small across all treatments.[26] Turning to the closeness centrality measure, we observe a steady increase over time and it reaches between 0.7 and 0.8 across all the treatments. Taking the data on average distances and closeness centrality together, we conclude that subjects create centralized networks with small average distances. It is worth noting that this pattern is very similar to what we observed in the Linking Game experiment.

Figures 10e and 11e show that the dominant status of the hub in an interval of 10 seconds is very likely to continue in the next time interval: around 80% in the small groups and strictly larger than 80% in the large groups, under both information conditions. The per capita number of linking choices per block of 10 seconds goes down to around 0.5 toward the end of the game across all the treatments. We therefore conclude that subjects create stable networks in which hubs persist and this persistence grows with group size.

These observations are summarized in the following statement.

Result 2.1 In the Connectors and Influencers experiment subjects create sparse, unequal, and small average distance networks, under all treatments. These findings on network structure are in line with Hypothesis 2A.

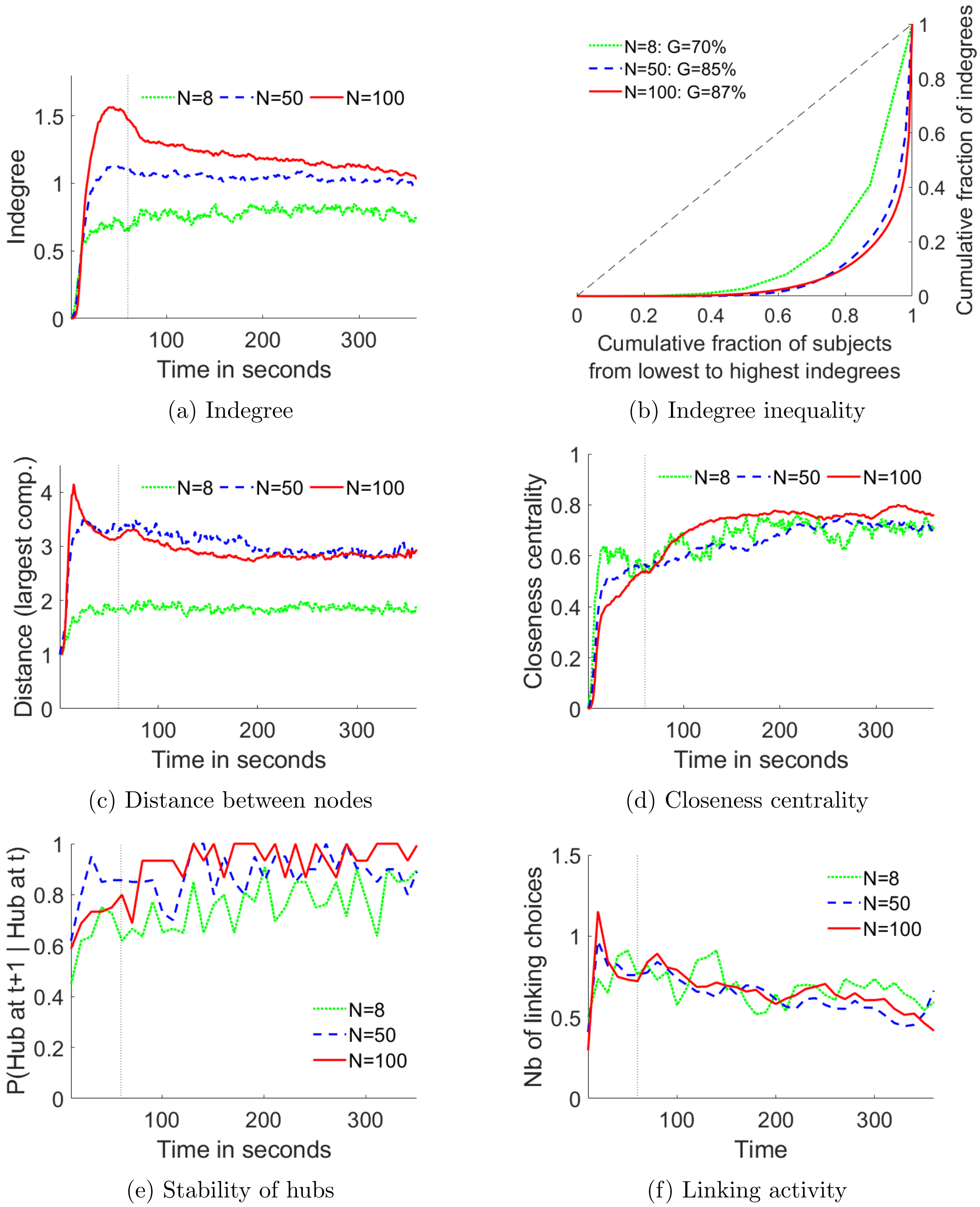
4.3.3 Efficiency

At any second of the round, we define efficiency as the ratio of observed total payoffs to the best Nash equilibrium payoff, i.e. the pure influencer outcome.

Figure 12 plots the time series of efficiency across group sizes for the two information treatments (the horizontal dashed line highlights the efficiency of the pure influencer outcome). In the baseline treatment, we observe positive group size effects on efficiency – it is around 0.95 for $N = 8$, above 1 for $N = 50$, and around 1.4 for $N = 100$ although there is some falling off toward the end of the game in the large groups ($p < 0.01$ from paired t-test comparing $N = 8$ or $N = 50$ with $N = 100$ with the group level average data). The reason why the level of efficiency in the large groups is higher than that of the best Nash equilibrium is that a few subjects make excessive effort investments in order to become hubs and this increases payoffs of all other subjects.

We do not observe any group size effects in the payoff information treatment (two-sample t-test with the group level average data: p=0.30 for $N = 8$ and $N = 50$, p=0.97

[26] The average size of the largest component is close to the group size in each treatment: 6.4 and 6.1 for Baseline8 and Payinfo8, 44.9 and 43.5 for Baseline50 and Payinfo50, and 94.8 and 93.5 for Baseline100 and Payinfo100.



(a) Indegree

(b) Indegree inequality

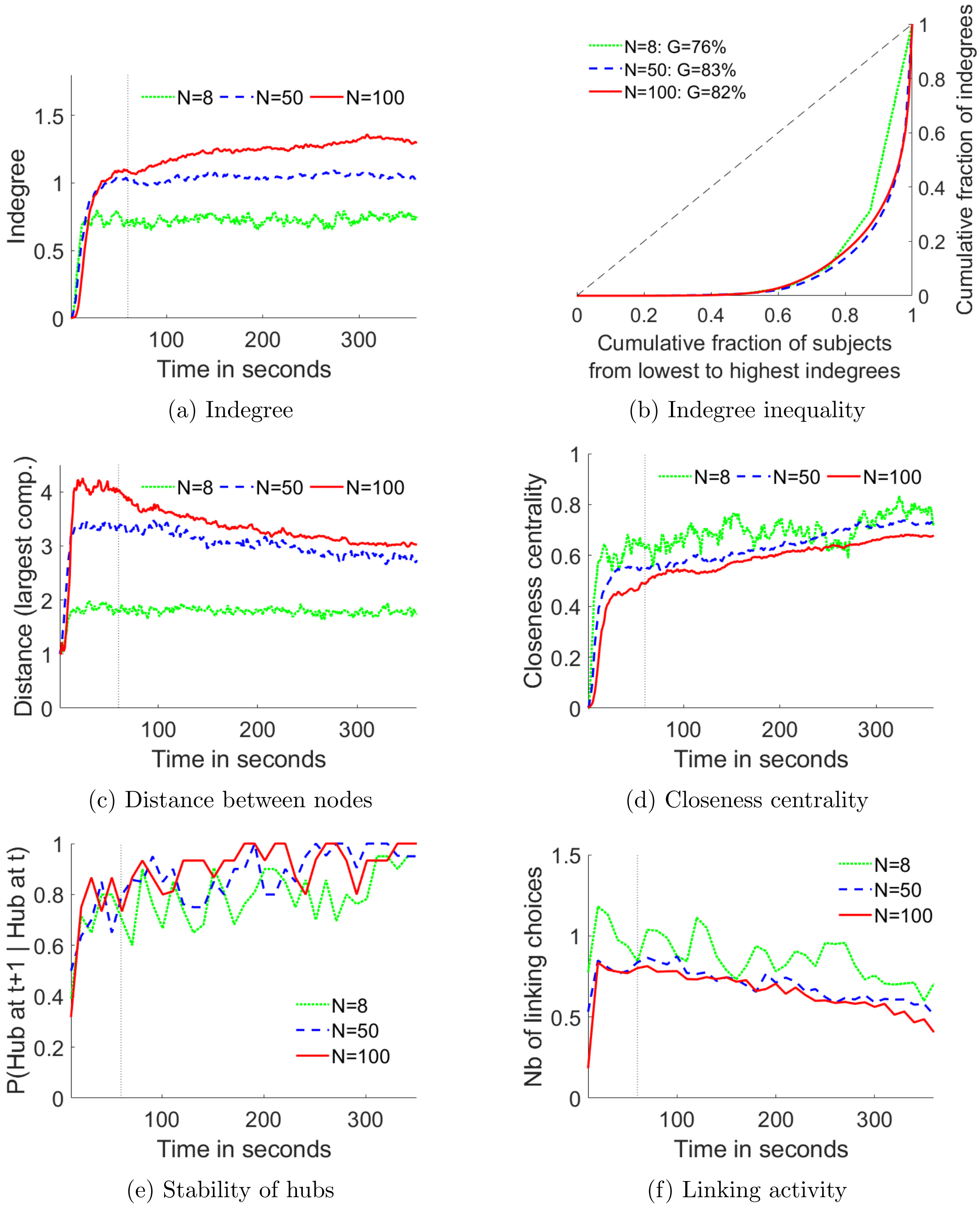
(c) Distance between nodes

(d) Closeness centrality

(e) Stability of hubs

(f) Linking activity

Figure 10: Network Structure in the Baseline Treatment



(a) Indegree

(b) Indegree inequality

(c) Distance between nodes

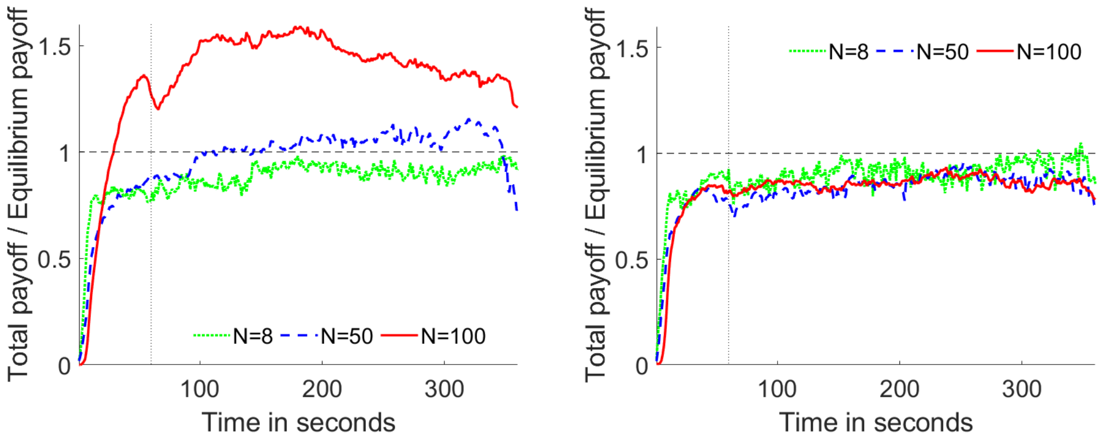
(d) Closeness centrality

(e) Stability of hubs

(f) Linking activity

Figure 11: Network Structure in the PayInfo Treatment

for $N = 50$ and $N = 100$, and p=0.30 for $N = 8$ and $N = 100$). The level of efficiency is slightly below 1.



(a) Baseline (b) PayInfo

Figure 12: Efficiency

We summarize these observations as follows.

Result 2.2 In the Connectors and Influencers experiments, subjects create networks and choose efforts that lead to payoffs that are close to or that exceed the best Nash equilibrium payoffs. In the baseline treatment, an increase in the group size leads to significantly larger payoffs than the best Nash equilibrium.

The group size effects are especially striking and lead to a closer examination of individual behavior.

4.3.4 Behavior and inequality

Figures 13 and 14 present the time series of efforts made and payoffs earned by the three different types of subjects – most connected, 2nd most connected, and the others.[27]

There are powerful effects of group size on the dynamics of efforts, in both information treatments. In the baseline treatment, in every group size, the two most connected subjects

[27]In Appendix G.2, we report the regression results of subject types – most connected, 2nd most connected, and others – on average behavior and median payoffs across different treatments. The effects of group size and information on efforts and payoffs across subjects' types, inferred from Figures 13 and 14, are supported by the regression analysis.

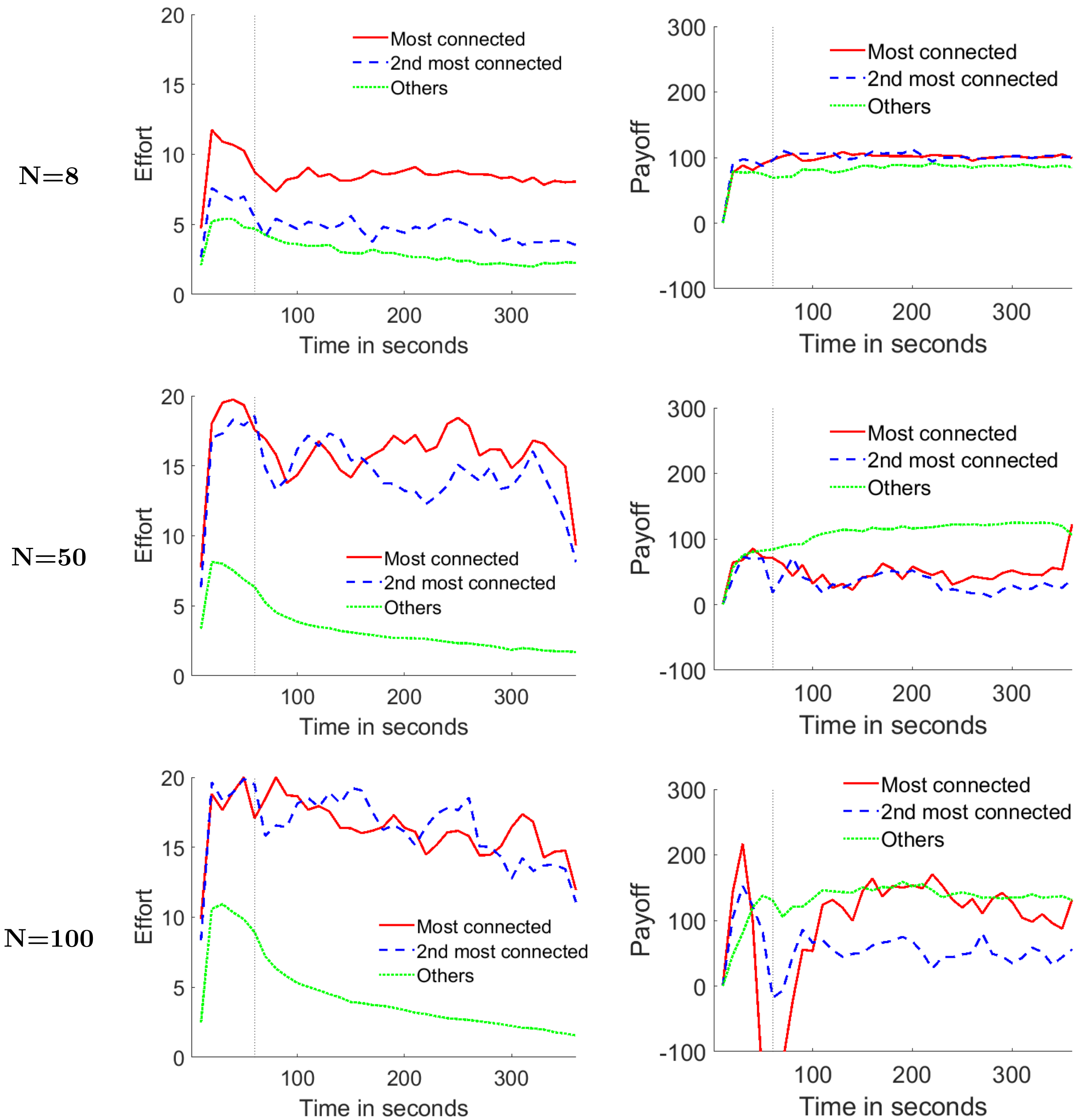


Figure 13: Behavior and inequality in baseline treatment.

compete for a hub position through efforts higher than those made by the other subjects. A

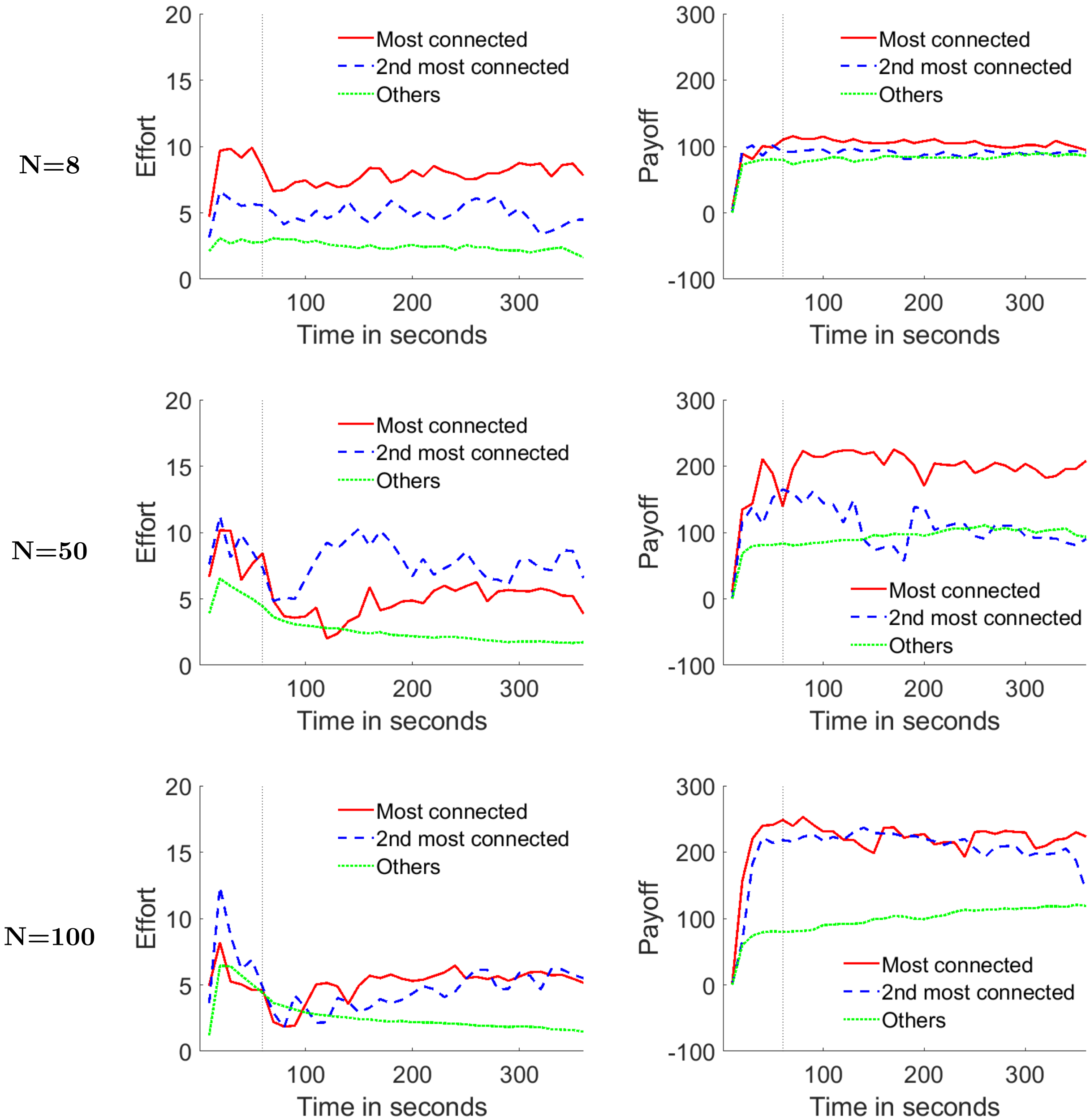


Figure 14: Behavior and inequality in the payoff information treatment.

remarkable feature of these dynamics is that this competition becomes much more intense as we increase the size of the groups: as a result, efforts made by the two most connected individuals in the large groups are much higher than the corresponding efforts in the small group. The behavior of 'other' subjects is however similar across different group sizes in the baseline treatment. Thus, the outcomes in the baseline treatment are in line with the pure influencer one, confirming Hypothesis 2B.

In the payoff treatment, the group size effects on efforts go in the opposite direction. The effort dynamics in the small group here are similar to those in the small group baseline treatment. However, efforts of the most connected individual in the large groups are much lower than those made by the most connected individual in the baseline treatment. As observed in the snap shots of Figure 9, on becoming the hub, the most connected individual shades her efforts substantially. There is however no effective challenge from her competitors and she retains her hub status. The behavior of 'other' subjects is similar across the two information treatment and across different group sizes. Hence, the outcomes in the large group payoff information treatment are close to the pure connector outcome, which is inconsistent with Hypothesis 2B.

These group size effects on efforts have serious implications for payoffs. In the baseline treatment, in the large groups, as a consequence of the intense competition, the two most connected individuals earn on average less than the other subjects in the same group and also less than the highly connected individuals in the small group. As a result, higher degree is associated with lower payoffs in the baseline large groups.

By contrast, in the payoff information treatment, the most connected individuals in the large groups earn substantially more than other subjects in the same group as well as the most connected individuals in the small group. As a result, higher degree is associated with higher payoffs.

We summarize these observations as follows:

Result 2.3 In the Connectors and Influencers experiment, group size interacts with payoff information and has powerful effects on linking and payoffs. In small groups, subjects coordinate on the pure influencer outcome under both information treatments. In the baseline treatment, an increase in group size leads to much greater efforts by the two most connected individuals and this leads to a lower payoff for them, relative to the other individuals. Subjects' choice of the pure influencer outcome is in line with Hypothesis 2B. In the payoff information treatment, an increase in group size leads to lower effort by the two most connected individuals and this leads to a higher payoff for them, relative to the other individuals. Subjects' choice of the pure connector outcome is inconsistent with Hypothesis 2B.

We conclude with a discussion of the strong interaction between scale and payoff information as this shapes effort levels of the hub and its competitor. This pattern of extreme competition is common to both our experiments – the Linking Game and the Connectors and Influencers. In the Linking Game and in the baseline treatment of the Connectors and Influencers experiment, these two individuals raise their investments. As a result, their earnings suffer and they earn significantly less than the other subjects. But in the payoff information treatment, the opposite happens: as we raise group size, these individuals actually lower effort and as a result higher payoffs as compared to the other subjects. The strong effects of payoff information on investment suggests that individuals in large groups who are making large investments in the baseline treatment do not appreciate that their strategy of large investments is not attractive as compared to other less active strategies. As these excessive investments happen only in large groups, it points to informational overload as an explanation for overinvestments.

5 Brokerage and Market Power

This section presents a model of intermediation rents and brokerage. There are two major differences in this experiment as compared to the two experiments presented above. The first difference is that linking is two-sided, in contrast to one-sided linking in the two experiments presented above. Two-sided linking calls for different design and network visualization (refer to the discussion in section 2). We note that two-sided linking considerably reduces the subjects' autonomy with regard to linking, as they need to coordinate with others to create links. The second major difference lies in the theoretical predictions: in the first two experiments, the prediction is a sparse, centralized and very unequal network, while in the third experiment, the theoretical prediction under one treatment will be a sparse, diffused and equal network.

5.1 Theory

We now consider a context in which the formation of links is more restrictive as it requires mutual consent between individuals. Specifically, players propose links with others, and those links are realized only if reciprocated. Formally, the strategy of a player i is a vector of link proposals $s_i = [s_{ij}]_{j \in N\setminus\{i\}}$, with $s_{ij} \in \{0,1\}$ for any $j \in N\setminus\{i\}$. The strategy set of player i is denoted by S_i. A link between agents i and j is formed if both propose a link to each other, i.e., $g_{ij} = s_{ij}s_{ji}$. A strategy profile $s = (s_1, s_2, \ldots, s_n)$ induces an undirected network $g(s)$.[28] There exists a path between i and j in a network g if either $g_{ij} = 1$, or if there is a distinct set of players $i_1, \ldots, i_n$ such that $g_{ii_1} = g_{i_1 i_2} = g_{i_2 i_3} = \ldots = g_{i_n j} = 1$. All players with whom i has a path defines the component of i in g, which is denoted by $C_i(g)$.

Suppose that players are traders who can exchange goods and that this exchange creates a surplus of V. This exchange can be carried out only if these traders have a link or if there is a path between them. There is a fixed cost k per individual for every link that is established. On the other hand, any proposal that is not reciprocated carries no cost.

In the case where two traders have a link, it is natural that they split the surplus equally, each earning $V/2$. If they are linked indirectly, then the allocation of the surplus depends on the nature of competition between the intermediary agents. One idea is to view these paths as being perfect substitutes. Another possibility is that the paths offer differentiated trading mechanisms.

Suppose that paths between traders are perfect substitutes. A trader is said to be *critical* for a pair of traders A and B if she lies on all paths between these traders. Trade between A and B occurs if there is a path between them. The surplus is divided equally between the traders and all critical intermediaries.[29] Denote by $T(j,k;g)$ the set of players who are critical for j and k in network g and let $t(j,k;g) = |T(j,k;g)|$. Following Goyal and Vega-Redondo [2007], for every strategy profile $s = (s_1, s_2, \ldots, s_n)$ the net payoffs to player i are given by:

$$\Pi_i^{crit}(s) = \underbrace{\sum_{j \in C_i(g)} \frac{V}{t(i,j;g)+2}}_{\text{Access Benefits}} + \underbrace{\sum_{j,k \in N\setminus\{i\}} V \frac{I_{i \in T(j,k;g)}}{t(j,k;g)+2}}_{\text{Brokerage Rents}} - \eta_i(g)k \tag{6}$$

where $I_{i \in T(j,k)} \in \{0,1\}$ stands for the indicator function specifying whether i is critical for j and k. We shall refer to it as the model of criticality-based pricing, which relies on two distinct sources of benefits: access benefits and brokerage rents.

[28] With a slight abuse of notation, for simplicity, we will write g instead of $g(s)$.

[29] Using a combination of theory and experiments, Choi, Galeotti, and Goyal [2017] show that, for any network, the surplus is divided (more or less) equally between the origin and destination traders and the critical traders, while the non-critical traders earn close to zero.

Let $n_{jk} = (d(j,k;g)-1)$ denote the number of intermediaries on a shortest path between j and k in network g. Trade surplus between j and k is equally distributed among the source and destination j and k, and among the intermediaries on the shortest path. In the case of multiple shortest paths, one of them is randomly chosen. Therefore, the (ex-ante) expected return for any trader i is in proportion to the shortest paths between j and k that i lies on. Formally, we write $b^i_{jk}(g) \in [0,1]$ to denote *betweenness* of player i between j and k.[30] Given a strategy profile $s = (s_1, s_2, \ldots, s_n)$, the net payoffs to player i are given by:

$$\Pi_i^{btwn}(s) = \underbrace{\sum_{j \in C_i(g)} \frac{V}{n_{ij}+2}}_{\text{Access Benefits}} + \underbrace{\sum_{j,k \in N \setminus \{i\}} V \frac{b^i_{jk}}{n_{jk}+2}}_{\text{Brokerage Rents}} - \eta_i(g)k \tag{7}$$

This is the model with betweenness pricing; we borrow this pricing rule from Kleinberg et al. [2008] and Galeotti and Goyal [2014].

We will study pairwise stable networks (see Jackson and Wolinsky [1996] for a formal definition). Goyal and Vega-Redondo [2007] establish the following result on network formation with criticality-based pricing.

Proposition 5.1. *Suppose payoffs are given by (6). There always exists a pairwise stable network. Pairwise stable networks include the empty network if* $k > \frac{V}{2}$*, the star network if* $\frac{V}{6} < k < \frac{Vn}{3} - \frac{V}{6}$*, and the cycle network if* $k < \sum_{i=1}^{n-2} \frac{Vi}{2(2+i)}$*. The complete network is not stable for* $n \geq 4$*.*

A general observation is that pairwise stable networks cover a wide range of structures that include the star and the cycle. So incentives in this model sustain networks with very small diameter as well as very large diameter. This also means that stability is not incompatible with efficiency or equality.

We next state a result on pairwise stable networks with betweenness-based pricing.

Proposition 5.2. *Suppose payoffs are given by (7). There always exists a pairwise stable network. Pairwise stable networks include the empty network if* $k > \frac{V}{2}$*, the complete network if* $k < \frac{V}{6}$*, and the star network if* $\frac{V}{6} < k < \frac{Vn}{3} - \frac{V}{6}$*. For a given value for* k *and* V*, the cycle is not pairwise stable for large* n*.*

[30] Formally, $b^i_{jk}(g) = \frac{\text{\# shortest paths between j and k on which i lies}}{\text{\# shortest paths between j and k}}$.

Pairwise stability of the empty network follows from noting that two isolated individuals hope to earn $V/2$ on forming a link. The conditions on the pairwise stability of the star network arise from two incentive constraints: spokes must not wish to form a link (this yields the constraint $V/6 < c$) and the central hub must wish to form a link with a spoke (this yields the constraint $c < Vn/3 - V/6$). In the complete network, no player can benefit by removing a link as long as $k < \frac{V}{6}$. In the cycle network, the gain in benefits (access benefits and brokerage rents) for adding a link between two players sitting at opposite points of the cycle increases with n. As a result, if n is sufficiently large, such a move becomes profitable for both players. Finally observe that for any values of $k > 0$ and $n \geq 3$, at least one of empty, star, and complete network is pairwise stable.

Turning to efficiency, observe that the intermediation rents cancel out when we sum across individuals. A network is said to be efficient if it maximizes the sum of trade surplus realized less the costs of links. As intermediation links cancel out, every component in an efficient network must be minimally connected or a singleton. Indeed, Goyal and Vega-Redondo [2007] prove that an efficient network is either the empty network or the minimally connected network. The total payoffs in the latter case are $\frac{Vn(n-1)}{2} - 2(n-1)k$ and they are equal to 0 in the case of an empty network. So it follows that an efficient network is minimally connected if $k < \frac{Vn}{4}$, and empty otherwise. A prominent example of minimally connected network is the star network.

Finally, payoff inequality significantly varies across different stable network structures. The outcome is equal in the empty network and in a cycle network. By contrast, in the star network (under both criticality and betwenness), the hub and spoke earn respectively:

$$V(n-1)\left[\frac{1}{2} + \frac{n-2}{6}\right] - (n-1)k \qquad\qquad V\left[\frac{1}{2} + \frac{n-2}{3}\right] - k \tag{8}$$

The ratio of the two payoffs grows without bound, in n, highlighting large inequalities in large groups.

Our experimental parameters are as follows. The value of trade between any two traders, $V = 10$. There are three groups sizes, 10, 50 and 100. The cost of a link is adjusted across scale to keep incentives as similar as possible. So the cost of a link is $k = 8$ for $n = 10$, $k = 40$ for $n = 50$, and $k = 80$ for $n = 100$.

We note that the star network is efficient in all three group sizes. It is however very unequal, especially as we raise the size of the group: the ratio of max payoff to median payoffs is 4 (for $n = 10$), 18 (for $n = 50$) and 35 (for $n = 100$). The cycle network by contrast is almost equally efficient and is perfectly equal. The differences in inequality between star and cycle are very large for large group sizes. This tension between inequality and efficiency is a key element in our design.

The analysis of pairwise stable networks, efficiency and inequality suggests the following hypotheses.

Hyopothesis 3A *Subjects create sparse networks for both pricing treatments and for all group sizes, as this is consistent with efficiency and individual incentives.*

Hyopothesis 3B *Under criticality pricing, subjects create equal and large average distance networks, as they reconcile efficiency with equity and individual incentives.*

Hyopothesis 3C *Under betweenness pricing, subjects create unequal and small average distance networks, as they reconcile efficiency with individual incentives.*

We conclude by noting that, as in the previous two experiments, there exists n different star networks, each corresponding to a different player being hub. In addition to this, the central individual has to coordinate on the links with each of the other individuals. Similarly, individuals need to coordinate links very finely in the cycle network so as to bring everyone within a cycle. Thus while the forces of efficiency and equity point to these networks, individuals face multiple challenges to arriving on such networks.

5.2 Experimental design and procedures

The experiment considers three group size – 10, 50 and 100 – and two pricing rules – criticality and betweenness. The experiment consists of 6 treatments in all: 3 group sizes $\times$ 2 pricing protocols.

In a round, at any moment, the subject is shown the entire network of reciprocated links. In addition, every subject is shown all outstanding link proposals – made and received – that involve them. Every subject is also provided full information on the payoffs of everyone (this is done by mentioning the numeric value of the payoffs for every subject next to their player ID. However, subjects are *not* shown unreciprocated links among other pairs. The principal motivation for this design choice was to keep the information options available to a subject manageable. Figure 28 presents the screen observed by subjects.

The number and duration of the rounds was as in the previous two experiments. Also, as in the earlier experiments, at any instant a subject can make or remove a proposal to another subject by simply double-clicking on the corresponding node in the computer screen. Any reciprocated proposal leads to the formation of a link. Non-reciprocated links were dealt with the protocol described in section 2. At any moment, every subject is shown the amount of access benefits, brokerage rents, overall cost of linking, and net payoffs. Finally, the subjects are also provided with information about the net payoffs of every other player (given within the corresponding node of the network). Figure 28 summarizes this information.

There were in total 18 sessions: 1 session with 4 groups of 10 subjects, 4 sessions with 50 subjects, and 4 sessions with 100 subjects for each of the Criticality and Betweenness treatments. A total of 1280 subjects participated in the experiment.

At the beginning of the experiment, every subject was endowed with an initial balance of 80 points for the $N = 10$ treatments, 400 points for the $N = 50$ treatments, and 800 points for the $N = 100$ treatments. Subjects' total earnings in the experiment were given by the sum of earnings across the last 5 rounds and the initial endowment. Earnings were calculated in terms of experimental points and then exchanged into euros at the rate of 20 points being equal to 1 euro in the $N = 10$ treatments, 110 points being equal to 1 euro in the $N = 50$ treatments, and 220 points being equal to 1 euro in the $N = 100$ treatments.[31] Each session lasted on average 90 minutes, and subjects earned on average about 16.4 euros (including a 5 euros show-up fee).

5.3 Results

5.3.1 Overview

We begin by presenting snap shots of the typical dynamics in groups with a hundred subjects. Figures 15a and 15b show the snap shots of the criticality treatment at minute 3 and minute 6, respectively. Network structures are sparse and connected and fairly dispersed. There is no single player who occupies a dominant network position and extracts large brokerage rents.

Figures 15c and 15d show that the dynamics in the betweenness treatment are quite different. At minute 3, one subject (represented in red) starts to emerge as a hub, and becomes a dominant hub at the end of the game. As a result, she earns substantial brokerage rents.

[31] As in the Linking Game experiment, different conversion rates and initial endowments are used to maintain similar earnings across treatments.

These snap shots draw attention to three points: first, under both pricing protocols, subjects create sparse and connected networks; two, the pricing protocol leads to the emergence of equal and dispersed networks under criticality and to unequal and small distance networks under betweenness pricing; third, there is little inequality in the criticality treatment while the hub in the betweenness treatment earns large brokerage rents and, as a result, there is great payoff inequality in the betweenness treatment. We now present a systematic analysis of the experimental data.

5.3.2 Network structure

Figures 16 and 17 present network properties observed across group sizes. pricing protocols.

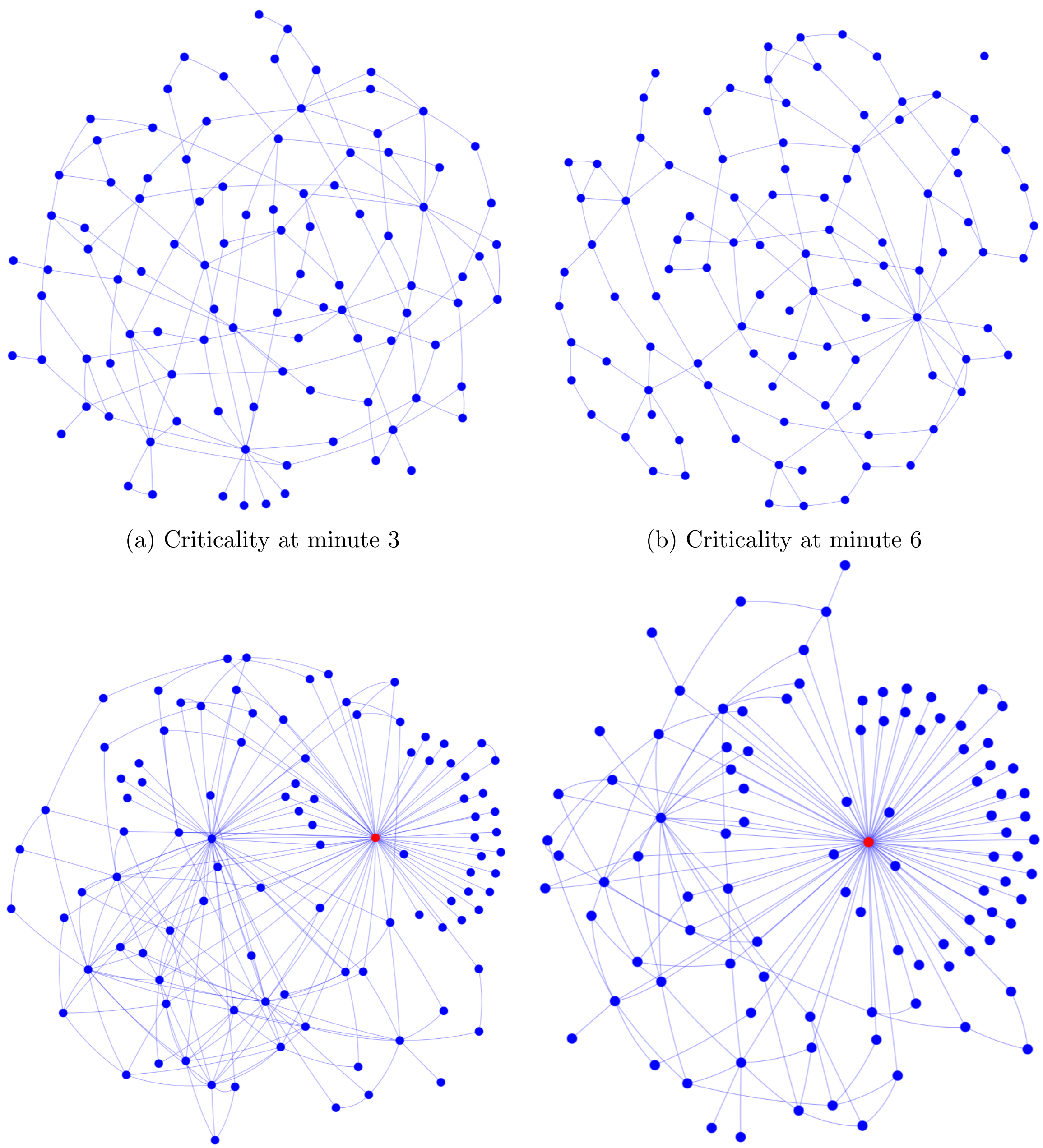
We start with the average degree. In the criticality treatment, the average degree lies between 2 and 3, across the different group sizes. In the betweenness treatment, the average degree is higher: stable around 3 for $N = 10$, falling from 5 to 3 for $N = 50$, and falling from 5 to 4 for $N = 100$. To conclude, average degree is low in all treatments, which suggests that subjects create sparse networks. This is consistent with Hypothesis 3A.

Second, consider inequality in linking. Figures 16b and 17b show that the Gini coefficients are 20% for $N = 10$, 26% for $N = 50$, 30% for $N = 100$ in the criticality treatment (two-sample t-test with the group level average data: $p < 0.01$ for $N = 10$ and $N \in \{50, 100\}$, and $p < 0.05$ for $N = 50$ and $N = 100$). The Gini coefficients are 29% for $N = 10$, 48% for $N = 50$, 51% for $N = 100$ in the betweenness treatment (two-sample t-test: $p < 0.01$ for $N = 10$ and $N \in \{50, 100\}$, and $p = 0.06$ for $N = 50$ and $N = 100$). Observe that in a two-sided link setting, the Gini coefficient of the star network is slightly below 50% across all group sizes.[32] We therefore conclude that degree inequality is very high in large groups under the betweenness treatment.

Third, we consider distance/closeness in networks. In the small group, the average distance (in the largest component) is around 2, under both pricing protocols. But size interacts strongly with the pricing protocol: in large groups, average distance is above 4 in the criticality treatment and below 3 in the betweenness treatment.[33] The difference between the pricing protocols is also salient when we consider closeness centrality. In the

[32]Indeed, the Gini coefficients for scale-free networks is 52% and 31% in Erdos-Renyi model with a similar average degree of 3.

[33]The average size of the largest component under the criticality pricing is 9.9 in the $N = 10$ group, 49.7 for the $N = 50$ group, and 99.3 for the $N = 100$. Under the betweenness pricing, it is 9.9 in the $N = 10$ group, 49.9 for the $N = 50$ group, and 99.1 for the $N = 100$



(a) Criticality at minute 3

(b) Criticality at minute 6

(c) Betweenness at minute 3

(d) Betweenness at minute 6

Figure 15: Snap shots from criticality and betweenness treatments

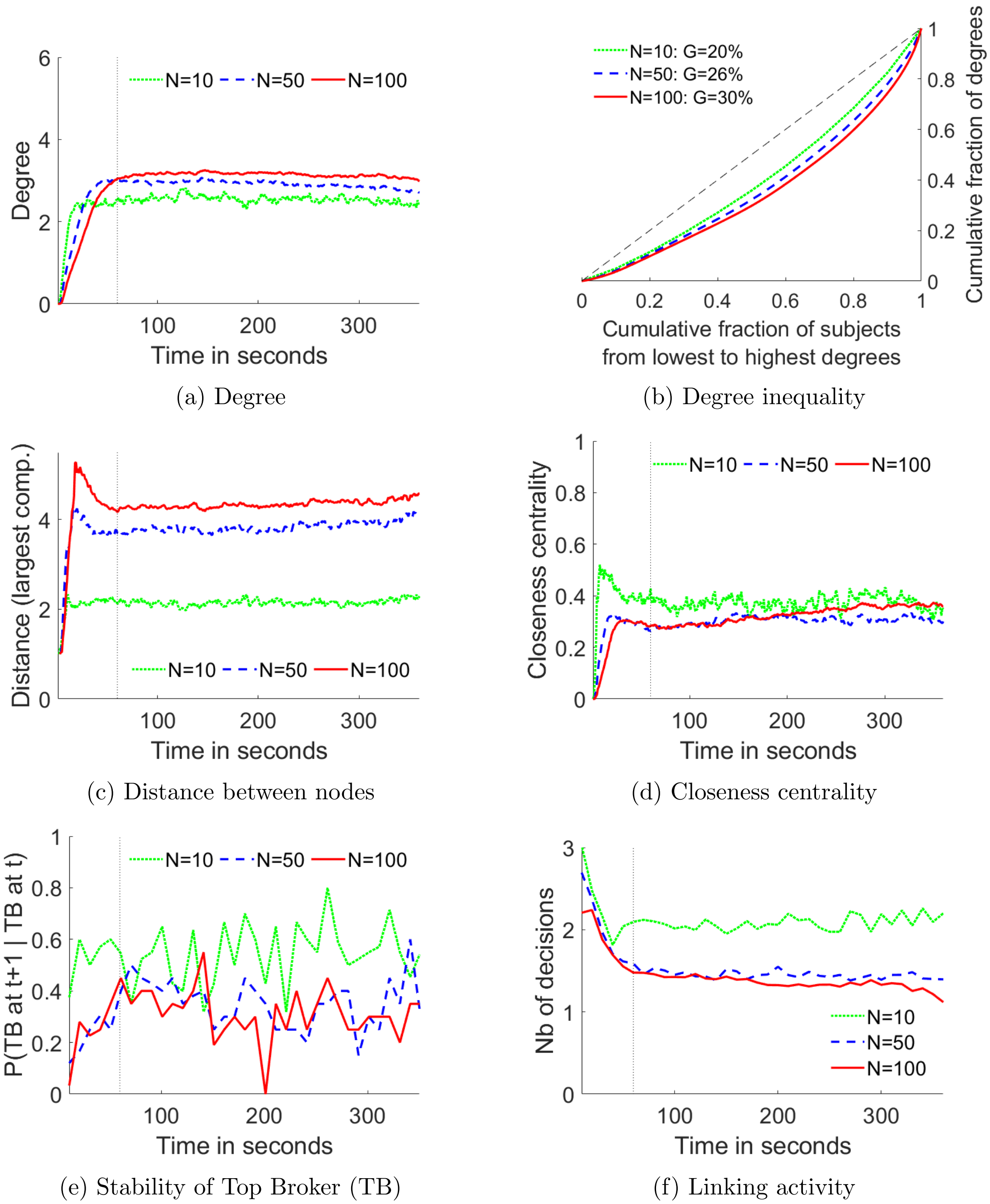
criticality treatment, subjects create networks with closeness centrality close to 0.4, in every group size. By contrast, subjects create networks with closeness centrality of 0.5 for $N = 10$ (two-sample t-test comparing pricing protocols: $p < 0.05$). This closeness centrality rises over time and is greater than 0.7 for the large groups (two-sample t-test comparing pricing protocols: $p < 0.01$). We conclude that pricing rules have a very strong effect on average distance and closeness centrality in large groups.

Under criticality pricing, subjects tend to create equal and spread out, large average distance networks. By contrast, under betweenness pricing, subjects create unequal and small average distance networks. Taken together, these findings on inequality of linking and distance/closeness support Hypothesis 3B and 3C.

Lastly, we observe powerful effects of pricing protocols on stability of networks. We focus on the persistence of the individual with the highest brokerage rents, because it crucially depends on network structure. For every interval of 10 seconds, we define the *Top Broker (TB)* as the subject who earns largest brokerage rents for the longest time within the interval. We study the likelihood of the top broker in a given interval remaining the top broker in the next interval of ten seconds. Figures 16e and 17e show that the status of top broker is highly persistent in the betweenness treatment, while it is quite unstable in the criticality treatment. This contrast is particularly strong in the large groups. To see this, fix N=100 group size: in the last 3 minutes of the game, the hub keeps her dominant status with near certainty under betweenness pricing, while she keeps this status with probability 0.40 in the criticality treatment. The networks become progressively stable under both pricing rules, as the per capita number of linking changes declines over time in large groups.

We summarize these observations as follows:

Result 3.1 In the Brokerage and Market Power experiment, subjects create sparse networks. Pricing protocols play a powerful role in shaping other features of networks: subjects create equal and spread out networks under criticality pricing and they create unequal and small average distance networks under betweenness pricing. Pricing protocols interact with group size so that these differences grow with scale. These findings are consistent with hypothesis 3A-3C.



(a) Degree

(b) Degree inequality

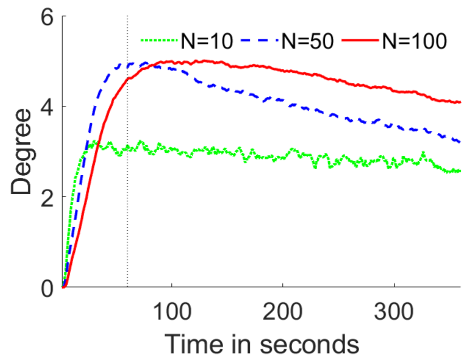
(c) Distance between nodes

(d) Closeness centrality

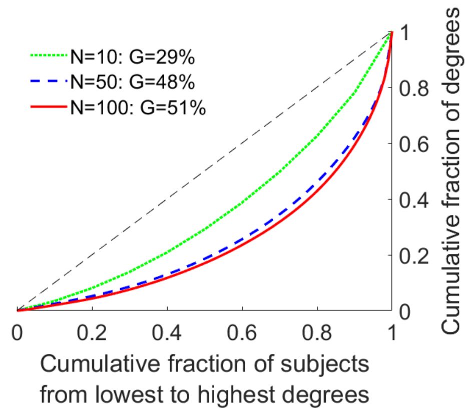
(e) Stability of Top Broker (TB)

(f) Linking activity

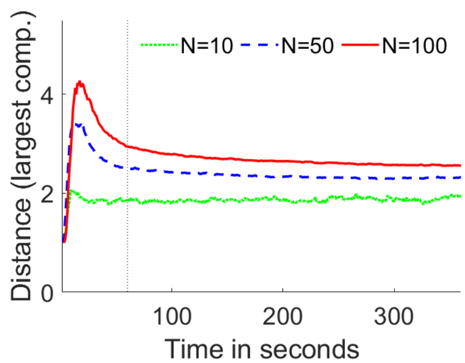
Figure 16: Network Structure in Criticality Treatment



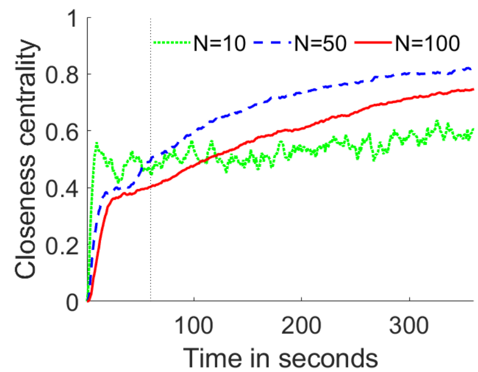
(a) Degree



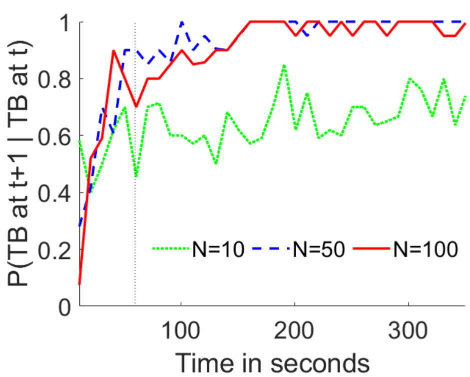
(b) Degree inequality



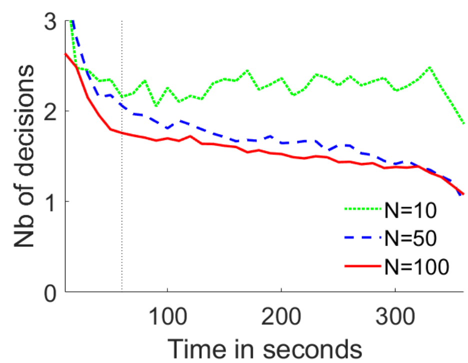
(c) Distance between nodes



(d) Closeness centrality



(e) Stability of Top Broker (TB)



(f) Linking activity

Figure 17: Network Structure in Betweenness Treatment

5.3.3 Efficiency

At any second, the level of efficiency is measured by the ratio of the total payoffs of an observed network, relative to the star network, i.e., the first-best pairwise stable and efficient network. The level of efficiency depends on (i) the connectivity of subjects (that is, the realization of trade between subjects) and (ii) the total number of links created by subjects. We note that the connectivity of subjects is very high and similar across treatments: on average, 98.6% for $N = 10$, 99% for $N = 50$, 98.7% for $N = 100$ for criticality pricing, and 97.5% for $N = 10$, 99.4% for $N = 50$, 98.1% for $N = 100$ for betweenness pricing.[34] Figure 18 plots the time series for efficiency levels.

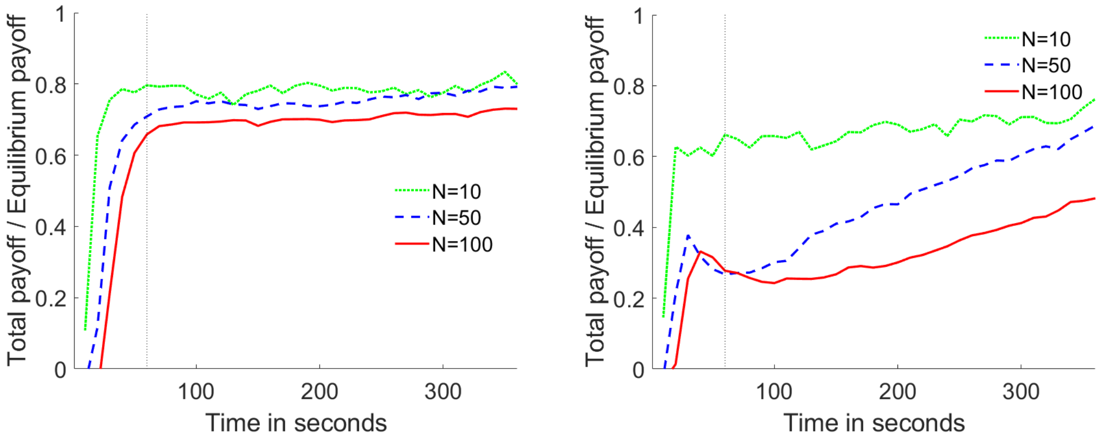
Pricing protocols have strong effects on efficiency. For any group size, efficiency is much higher in the criticality treatment than in the betweenness treatment. For group size $N = 10$, efficiency is around 0.8 and stable over time under criticality pricing, and it steadily increases to reach around 0.7 under betweenness pricing. The differences are more significant for larger groups: it is around 0.7 under criticality pricing, while it starts very low but then increases rapidly to reach 0.6 for $N = 50$ and over 0.4 for $N = 100$ under betweenness pricing (two-sample t-tests comparing average efficiency at the group level across pricing protocols, given any group size: $p < 0.01$).

Group size also appears to have effects on efficiency. There are minor negative effects of group size on efficiency under criticality pricing (two-sample t-tests comparing average efficiency at the group level for different group sizes: $p < 0.01$ for $N = 10$ and $N = 100$; $p < 0.05$ for $N = 50$ and $N \in \{10, 100\}$), but rather large group size effects on efficiency under betweenness pricing (two-sample t-tests comparing average efficiency at the group level across any group sizes: $p < 0.01$).

Given that the connectivity of subjects is very high and similar across treatments, i.e. that there is little loss of efficiency due to breakdown of trade, the main source of the differences of efficiency across treatments lies in the variations in the number of links. Figures 16a and 17a show the effects of group size and pricing protocol on the number of links created by subjects. The differences are large and they help explain the differences in efficiency across group sizes and pricing protocols. Therefore we draw the conclusion that over-linking is the principal source of inefficiency in the betweenness pricing.

We summarize the findings on efficiency as follows:

[34] Here connectivity is determined as the fraction of pairs of players that are directly or indirectly connected in the network (i.e., there exists a path connecting them in the network).



(a) Criticality (b) Betweenness

Figure 18: Payoff Efficiency

Result 3.2 In the Brokerage and Market Power experiment, both pricing rules and group size affect efficiency. For every group size, betweenness pricing leads to a lower level of efficiency. An increase in group size reduces efficiency moderately under criticality, but lowers it substantially under betweenness pricing. This loss in efficiency loss is a consequence of over-linking.

The efficiency attained in the 100 subjects group – under betweenness is an outlier to all the other experiments in this paper. A possible explanation may proceed along the following lines. A link under criticality pricing has large effects that are invariant with respect to the length of the path created. By contrast, an individual who contemplates a link under betweenness pricing, needs to consider the effects on the length of the different shortest paths. In a group with 100 subjects, this difference in informational requirements appears to be important.

5.3.4 Behavior and inequality

We study the number of link proposals made by the three different types of subjects (measured in terms of how many link proposals they have received): the most popular individual, the 2nd most popular individual, and the other individuals. Figures 19 and 20 plot the time series of the average fraction of the number of link proposals made by each type to the total number of link proposals and the median payoffs earned.[35]

In the large groups, there are major differences in the link proposals made by the two most popular individuals between the pricing protocols. In the betweenness treatment, a few individuals compete for the hub position by making a large number of link proposals; we observe no such competition in the criticality treatment. Notably, in the betweenness treatment, the fraction of link proposals by the most popular individual is growing over time. As a result, the most popular individual strengthens her status as hub, over time. This contrast between the two pricing protocols is less visible in the small group.

These differences in behavior translate into large differences in payoffs. Under criticality pricing, there is little inequality of payoffs, in every group size. By contrast, the hub (and most popular) individual earns vastly higher payoffs as compared to the rest of the subjects, in the betweenness treatment. Thus payoff inequality explodes, as the group size increases. For instance, the payoffs of the most popular individual reach 1500 in a group of size $n = 50$ and 5000 for $n = 100$, while the other subjects earn rarely more than 70 and 100 respectively.

What is the source of such a large inequality in payoffs? To address this question, we present the distribution of brokerage rents. Specifically, we first compute for each round the average payoffs earned by subjects over the last 5 minutes of the game and rank them from the lowest to the highest. We then take, for each rank, an average of brokerage rents over the last 5 minutes of the game. We normalize each subject's brokerage rents by the maximal brokerage rents that could be obtained by an individual in the group (150 in $n = 10$, 4083 in $n = 50$, and 16500 in $n = 100$). Figure 21 presents the bar graphs averaged across rounds and groups of the resulting distribution of brokerage rents in each treatment. There is little inequality of brokerage rents under criticality pricing. In contrast, we observe large inequality of brokerage rents across individuals for every group size under betweenness pricing. This inequality becomes much more salient in the large groups than in the small group. We conclude that the main source of payoff inequality is the extremely unequal distribution of brokerage rents.

We summarize the findings as follow:

Result 3.3 In the Brokerage and Market Power experiment, scale interacts with pricing rules and has large effects on behavior and payoffs. Criticality pricing gives rise to

[35] In Appendix G.3, we report the regression results of subject types – most popular, 2nd most popular, and others – on average outdegree and median payoffs across different treatments. The regression results provide statistical support to the patterns made from Figures 19 and 20.

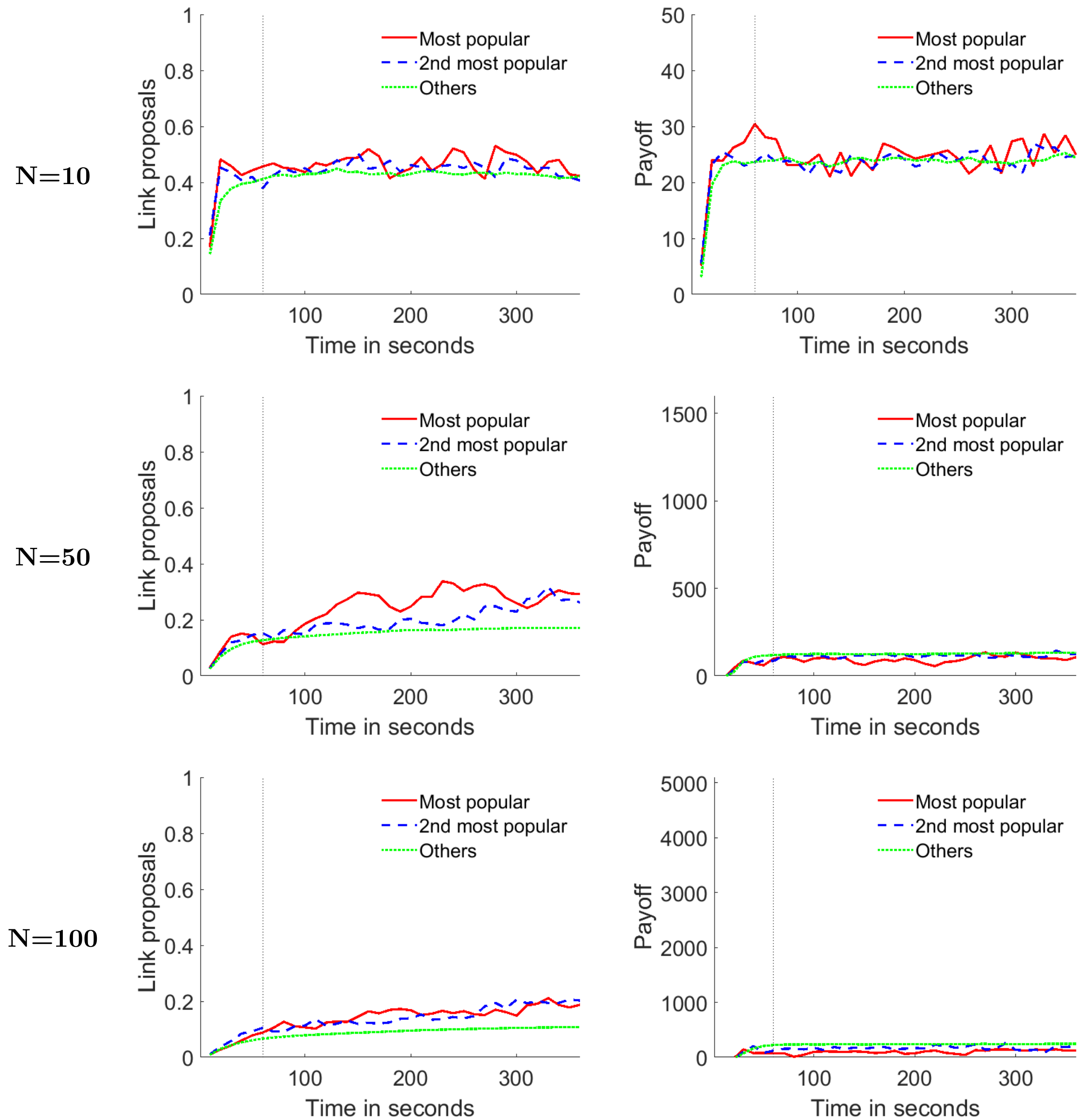


Figure 19: Behavior and inequality under criticality.

few link proposals and equal networks that yield similar payoffs to all subjects, in

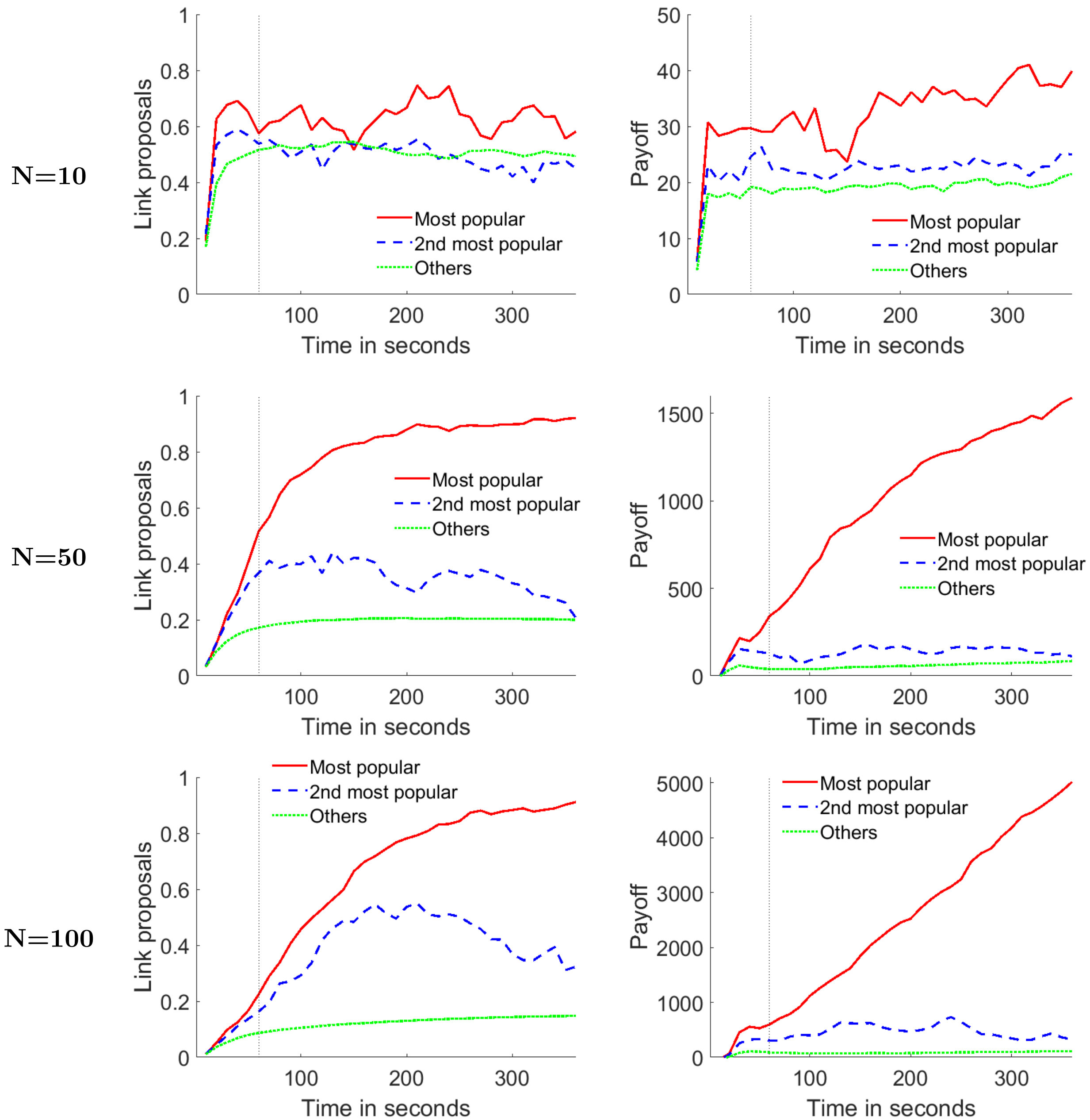
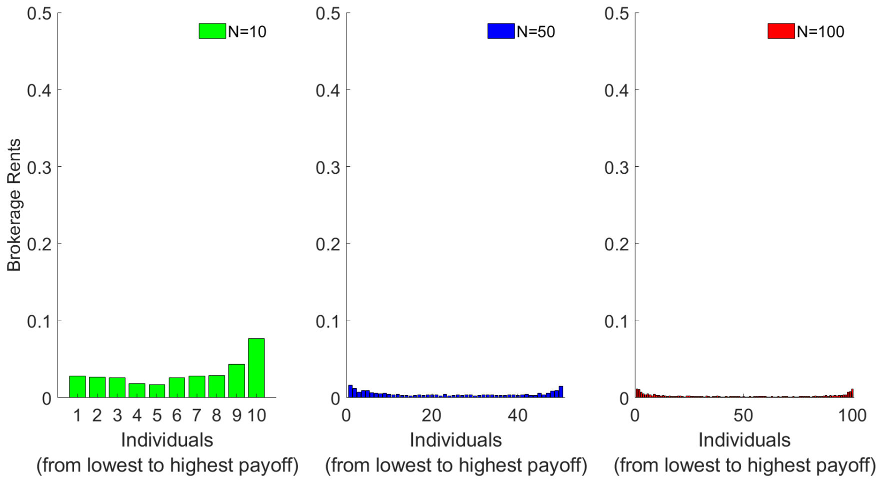
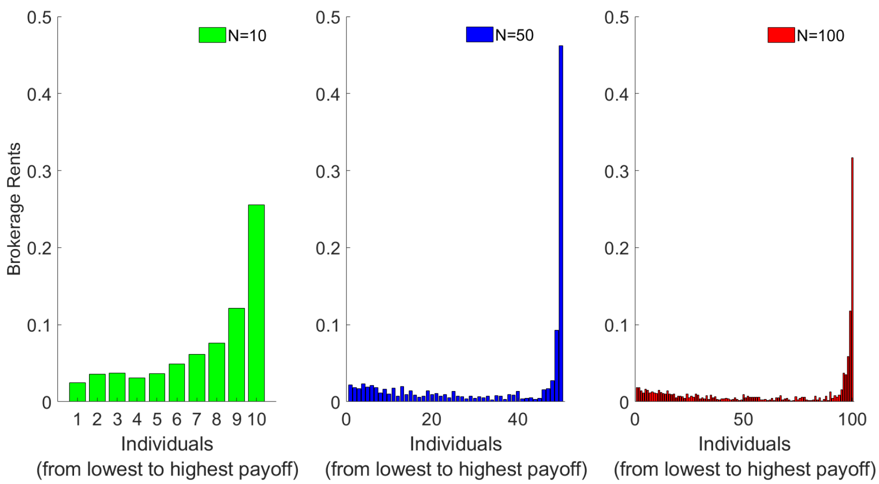


Figure 20: Behavior and inequality in Betweenness.

all group sizes. Betweenness pricing leads to a few subjects proposing many links



(a) Criticality



(b) Betweenness

Figure 21: Distribution of Brokerage Rents

and the emergence of a dominant hub with large brokerage payoffs, yielding highly unequal payoffs. The contrast between criticality pricing and betweenness pricing grows massively with group size.

6 Conclusion

Social, economic and infrastructure networks are an important feature of an economy. The economic theory of networks helps us in developing an understanding of how networks shape behavior and aggregate outcomes. At the heart of these models is purposeful behavior by rational and self interested individuals. Many real world networks are large: informational and computational complexity grows with scale as does inequality. These considerations are known to affect individual behavior. To appreciate the scope of these theories, it is therefore imperative that we conduct experiments with larger groups. This paper presents a new platform that allows us to conduct experiments with up to 100 subjects. A distinctive feature of the platform is that choice is asynchronous and takes place in continuous time – this in turn calls for a number of innovations in software and methodology.

This paper presents three experiments conducted on this platform. Group sizes range all the way from 8 to 100 subjects. These experiments involve linking and assorted activities.

In our experiments, subjects create sparse networks that are almost always highly efficient. In some experiments the networks are centralized, unequal, and have short average distances, while in others they are dispersed, equal, and have long average distances. These network structures are in line with theoretical predictions, suggesting that continuous time asynchronous choice facilitates a good match between experimental outcomes and theory. The second finding is that scale has powerful effects on network structure, on individual investments in linking and effort, and on the nature of payoff inequality. Researchers should therefore exercise caution in drawing inferences about behaviour in large scale networks based on data from small group experiments.

These experiments bring out some dimensions of the versatility of the platform. In ongoing work we are using the platform to conduct experiments on information aggregation and games on large scale networks. There are other applications, including propagation of shocks in an economy, in which scale and real-time interaction can matter. Our platform provides a new tool with which the validity of small-scale experiments is tested and causal inferences on large-scale phenomena are established.

References

M. Agranov and M. Elliott. Commitment and (in)efficiency: A bargaining experiment. *Journal of European Economic Association, forthcoming*, 2020.

J. Andreoni and J. Miller. Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica*, 70(2):737–753, 2002.

O. Baetz. Social activity and network formation. *Theoretical Economics*, 10:315–340, 2015.

V. Bala and S. Goyal. A non-cooperative model of network formation. *Econometrica*, 68 (5):1181–1231, 2000.

A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286 (5439):509–512, 1999.

J. Barnes and P. Hut. A hierarchical o(n log n) force-calculation algorithm. *Nature*, 324: 446–449, 1986.

S. K. Berninghaus, K.-M. Ehrhart, and M. Ott. A network experiment in continuous time: The influence of link costs. *Experimental Economics*, 9(3):237–251, 2006.

G. E. Bolton and A. Ockenfels. Erc: A theory of equity, reciprocity, and competition. *American economic review*, 90(1):166–193, 2000.

M. Bostock, V. Ogievetsky, and J. Heer. D^3 data-driven documents. *IEEE transactions on visualization and computer graphics*, 17(12):2301–2309, 2011.

Y. Bramoulle and R. Kranton. Public good in neworks. *Journal of Economic Theory*, 135 (1):478–494, 2007.

Y. Bramoulle, A. Galeotti, and B. Rogers. *Oxford Handbook on the Economics of Networks*. Oxford University Press, 2016.

E. Breza. Field experiments, social networks and development. In Y. Bramoulle, A. Galeotti, and B. Rogers, editors, *Oxford Handbook of the Economics of Networks*. Oxford University Press, 2016.

E. Calford and R. Oprea. Continuity, inertia, and strategic uncertainty: A test of the theory of continuous time games. *Econometrica*, 85:915–935, 2017.

S. Callander and C. R. Plott. Principles of network development and evolution: an experimental study. *Journal of Public Economics*, 89(8):1469–1495, 2005.

C. Camerer. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press, 2003.

C. Camerer and T. H. Ho. Experience-weighted attraction learning in normal form games. *Econometrica*, 67:827–874, 1999.

C. Camerer, T. H. Ho, and J.-K. Chong. Sophisticated experience-weighted attraction learning and strategic teaching in repeated games. *Journal of Economic Theory*, 104: 137–188, 2002.

D. Centola and A. Baronchelli. The spontaneous emergence of conventions: An experimental study of cultural evolution. *Proceedings of the National Academy of Sciences*, 112(7):1989–1994, 2015.

A. G. Chandrasekhar, H. Larreguy, and J. P. Xandri. Testing models of social learning on networks: Evidence from two experiments. *Econometrica*, 2019.

G. Charness, M. Corominas-Bosch, and G. R. Frechette. Bargaining and network structure: An experiment. *Journal of Economic Theory*, 136(1):28–65, 2007.

G. Charness, F. Feri, M. A. Meléndez-Jiménez, and M. Sutter. Experimental games on networks: Underpinnings of behavior and equilibrium selection. *Econometrica*, 82(5): 1615–1670, 2014.

G. B. Charness and M. Rabin. Understanding social preferences with simple tests. *Quarterly Journal of Economics*, 117:817–869, 2002.

S. Choi, E. Gallo, and S. Kariv. Networks in the laboratory. In Y. Bramoulle, A. Galeotti, and B. Rogers, editors, *Oxford Handbook of the Economics of Networks*. Oxford University Press, 2016.

S. Choi, A. Galeotti, and S. Goyal. Trading in networks: theory and experiments. *Journal of the European Economic Association*, 15(4):784–817, 2017.

D. Condorelli and A. Galeotti. Intermediation in networks. In Y. Bramoulle, A. Galeotti, and B. Rogers, editors, *Oxford Handbook of the Economics of Networks*. Oxford University Press, 2016.

D. Condorelli, A. Galeotti, and L. Renou. Bilateral trading in networks. *The Review of Economic Studies*, 84(1):82–105, 2017.

P. Eades. A heuristic for graph drawing. *Congressus numerantium*, 42:149–160, 1984.

P. Erdos and A. Rényi. On random graphs i. *Publ. math. debrecen*, 6(290-297):18, 1959.

A. Falk and M. Kosfeld. Its all about connections: Evidence on network formation. *Review of Network Economics*, 11, 2012.

E. Fehr and K. M. Schmidt. A theory of fairness, competition, and cooperation. *The quarterly journal of economics*, 114(3):817–868, 1999.

F. Ferri. Stochastic stability in networks with decay. *Journal of Economic Theory*, 135: 442–457, 2007.

D. Friedman and R. Oprea. A continuous dilemma. *American Economic Review*, 102(1): 337–63, 2012.

T. M. Fruchterman and E. M. Reingold. Graph drawing by force-directed placement. *Software: Practice and experience*, 21(11):1129–1164, 1991.

D. M. Gale and S. Kariv. Trading in networks: A normal form game experiment. *American Economic Journal: Microeconomics*, 1(2):114–32, 2009.

A. Galeotti and S. Goyal. The law of the few. *The American Economic Review*, 100(4): 1468–1492, 2010.

A. Galeotti and S. Goyal. Competing supply chains. *Mimeo. Cambridge and Essex*, 2014.

J. K. Goeree, A. Riedl, and A. Ule. In search of stars: Network formation among heterogeneous agents. *Games and Economic Behavior*, 67(2):445–466, 2009.

S. Goyal and F. Vega-Redondo. Structural holes in social networks. *Journal of Economic Theory*, 137(1):460–492, 2007.

S. Goyal, S. Rosenkranz, U. Weitzel, and V. Buskens. Information acquisition and exchange in social networks. *Economic Journal*, 127:2302–2331, 2017.

B. Herskovic and J. Ramos. Acquiring information through peers. *American Economic Review*, 110:2128–52, 2020.

D. Hojman and A. Szeidl. Core and periphery in networks. *Journal of Economic Theory*, 139:295–309, 2008.

C. A. Holt and S. K. Laury. Risk aversion and incentive effects. *American Economic Review*, 92(5):1644–1655, 2002.

Y. Hu. Efficient, high-quality force-directed graph drawing. *Mathematica Journal*, 10(1): 37–71, 2005.

S. Huck, H.-T. Normann, and J. Oechssler. Learning in cournot oligopoly-an experiment. *Economic Journal*, 109:C80–C95, 1999.

R. M. Isaac and J. M. Walker. Group size effects in public goods provision: The voluntary contributions mechanism. *The Quarterly Journal of Economics*, 103(1):179–199, 1988.

M. Jackson and A. Wolinsky. A strategic model of social and economic networks. *Journal of Economic Theory*, 71:44–74, 1996.

M. Jacomy, T. Venturini, S. Heymann, and M. Bastian. Forceatlas2, a continuous graph layout algorithm for handy network visualization designed for the gephi software. *PloS one*, 9(6), 2014.

J. H. Kagel and D. Levin. The winner's curse and public information in common value auctions. *The American economic review*, pages 894–920, 1986.

S. Kariv, M. H. Kotowski, and C. M. Leister. Liquidity risk in sequential trading networks. *Games and Economic Behavior*, 109:565–581, 2018.

M. Kearns, S. Judd, and Y. Vorobeychik. Behavioral experiments on a network formation game. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 690–704, 2012.

J. Kleinberg, S. Suri, É. Tardos, and T. Wexler. Strategic network formation with structural holes. In *Proceedings of the 9th ACM Conference on Electronic Commerce*, pages 284–293. ACM, 2008.

M. Kosfeld, A. Okada, and A. Riedl. Institution formation in public goods games. *American Economic Review*, 99(4):1335–55, 2009.

S. Leider, M. Mobius, T. Rosenblat, and Q.-A. Do. Directed altruism and enforced reciprocity in social network. *Quarterly Journal of Economics*, 124(1):1815–1851, 2009.

M. Manea. Intermediation and resale in networks. *Journal of Political Economy*, 126(3): 1250–1301, 2018.

A. Mauleon and V. Vannetelbosch. Network formation games. In Y. Bramoulle, A. Galeotti, and B. Rogers, editors, *Oxford Handbook of the Economics of Networks*. Oxford University Press, 2016.

J. Perego and S. Yuksel. Searching for information. *Working Paper Columbia University*, 2016.

J. Pettit, D. Friedman, C. Kephart, and R. Oprea. Software for continous game experiments. *Experimental Economics*, 17:631–648, 2014.

B. Rammstedt and O. P. John. Measuring personality in one minute or less: A 10-item short version of the big five inventory in english and german. *Journal of research in Personality*, 41(1):203–212, 2007.

A. Riedl, I. M. Rohde, and M. Strobel. Efficient coordination in weakest-link games. *The Review of Economic Studies*, 83(2):737–767, 2016.

A. Rubinstein and A. Wolinsky. Middlemen. *The Quarterly Journal of Economics*, 102(3): 581–593, 1987.

K. Schlag. Why imitate, and if so, how? a bounded rational approach to multi-armed bandits. *Journal of Economic Theory*, 78:130–156, 1998.

H. Schumacher, I. Kesternich, M. Kosfeld, and J. Winter. One, two, many—insensitivity to group size in games with concentrated benefits and dispersed costs. *The Review of Economic Studies*, 84(3):1346–1377, 2017.

B. van Leeuwen, T. Offerman, and A. Schram. Competition for status creates superstars: An experiment on public good provision and network formation. *Journal of the European Economic Association*, 18:666–707, 2020.

ONLINE APPENDICES

A Theory

Proof of Proposition 1 The first step is to observe that in equilibrium, every individual must access at least $\hat{y}$. This is true because if someone is accessing less than $\hat{y}$, then due to the concavity of the $f(.)$ function, she can simply increase her utility by raising effort so that the total access equals $\hat{y}$.

The second step is to show that players will form one link or zero link, for sufficiently large linking costs. Observe that an isolated individual will choose $\hat{y}$. So it follows that in a network with connections, no one will ever choose more than $\hat{y}$. Note that if link costs are close to $c\hat{y}$ then it is not profitable to form links with two individuals who each chooses $\hat{y}$. So the only situation in which an individual, A, may choose two or more links arises if an individual accesses significantly more than $\hat{y}$ through each link. Consider a link between A and B. Iterating on optimal effort, it is true that if B chooses $\hat{y}$ then every neighbor of B must choose 0. So A accesses more than $\hat{y}$ only if B chooses strictly less than $\hat{y}$. If a neighbour of B chooses a positive effort, then it must be the case that this person must meet the first order condition on optimal efforts: her total efforts invested and accessed must equal $\hat{y}$. As this person is a neighbour of B, it follows that A cannot access more than $\hat{y}$ via the link with B. So, A will form at most one link in equilibrium.

The third step considers effort configurations. Take the situation in which some individual (say) A chooses $\hat{y}$. It is optimal for everyone else to choose effort 0 and form a link with this person. And it is clearly optimal for A to choose $\hat{y}$ when faced with zero efforts by everyone else.

To conclude the proof, we need to show that the pure connector outcome is the only possible equilibrium in a situation where no player chooses $\hat{y}$. Observe first that the pure connector outcome is an equilibrium so long as $k < c\hat{y}(n-2)a_2/(1+(n-2)a_2)$. Observe that $c\hat{y}(n-2)a_2/(1+(n-2)a_2)$ converges to $c\hat{y}$, as n gets large.

The next step is to rule out any other possible equilibrium. The key observation here is that any equilibrium network must have diameter less than or equal to 2. Suppose the diameter of a component is 3 or more. We know from step 2 that the component must be acyclic. So consider two furthest apart leaf nodes. A variant of the 'switching' argument, developed in Bala and Goyal [2000], shows that one of the two leaf players has a strict incentive to deviate. So every component must have diameter 2. Given that the network is acyclic, this implies it must be a star. It is now possible to apply standard agglomeration arguments to deduce that multiple components cannot be sustained in equilibrium.

Finally, the hub player must choose zero. Suppose not. By hypothesis the hub chooses less than $\hat{y}$. Given that a_1 and $a_2 < 1$, both the hub and the spokes cannot be accessing exactly $\hat{y}$. A contradiction that implies that the hub must choose zero effort.

B Network visualization

The use of force-directed algorithms to improve the quality of graph visualization is not new to the literature. Any such algorithm relies on formulas that simulate both attraction and repulsion forces between nodes in the network. A standard model is the *"Spring-Electric"* layout [Eades, 1984], which is inspired by real physical phenomena: it uses the repulsion formula of electrically charged particles and the attraction formula of springs involving the geometric distance between two nodes. This specific model, combined with an additional gravity formula, is embedded in the *vis.js* open source library used in our software.[36] This algorithm has been used to conduct experiments 1 (A Linking Game) and 2 (Connector and Influencers) presented in this paper. However, other alternative methods have been proposed, relying on different formulas to model such forces [see, e.g., Fruchterman and Reingold, 1991, Hu, 2005, Bostock et al., 2011, Jacomy et al., 2014]. In particular, the *"ForcedAtlas2"* algorithm introduced by Jacomy et al. [2014] (also available in the *vis.js* library in a slightly modified version) has shown to be relevant in the context of social networks. Indeed, a common feature of this type of network is the presence of many "spokes" (nodes that have only one neighbor), which is due to the power-law distribution of degrees that characterizes many real-world data. The algorithm takes into account the degree of the nodes in the repulsion formula, which aims at reducing the visual cluttering that may otherwise be caused by the large number of spokes surrounding few highly connected nodes. Given its desirable properties in terms of performance and its relevance to visualize complex social networks, we have used this algorithm in the context of experiment 3 (Brokerage and Market Power). We note however that under

[36] We note the existence of other available open source libraries that use a similar approach to visualize complex graphs. Examples include the *D3* visualization tool by Bostock et al. [2011], and *Gephi* by Jacomy et al. [2014]. The choice of *vis.js* is justified by the javascript format (suitable for web based programming) and the wide range of parameter settings allowing rich interactions with the networks.

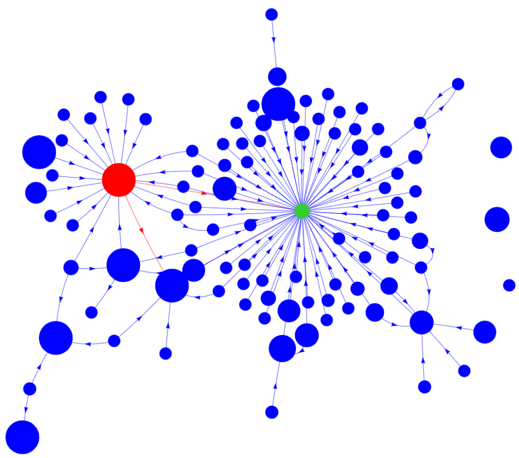
our specific choice of parameters, the *Spring-Electric* model would not offer significantly worse visualization given the networks actually observed in this experiment. Looking at Figure 22, which compares the visualization of networks observed in experiments 2 and 3 according to both algorithms, we observe that *ForcedAtlas2* offers a clearer visibility of sparse networks with an homogeneous distribution of degrees, whereas the *Spring-Electric* model would instead provide a more dense representation (see Figures 22c and 22d).

Beyond improving the quality of network visualization, another important constraint of all force-directed algorithms lies in their practicality to be computed in continuous time. In fact, those algorithms can be associated with large time complexity in the context of large networks. In particular, since repulsion forces are applied between any pair of nodes in the network, the time complexity of the corresponding algorithms is $O(n^2)$ (with n being the group size). To address this limitation, the *vis.js* package used in our software relies on the well-known Barnes-Hut algorithm [Barnes and Hut, 1986] to compute repulsion forces, which considerably reduces the time complexity to $O(nLog(n))$. Although the limited network sizes considered in our experiments (no more than 100 nodes) would not necessitate such approximations per se (this approximation becomes particularly relevant to visualize much larger networks with tens of thousands of nodes), the combination of other computationally demanding tasks justifies its use (e.g., payoff calculation is particularly complex in experiment 3).

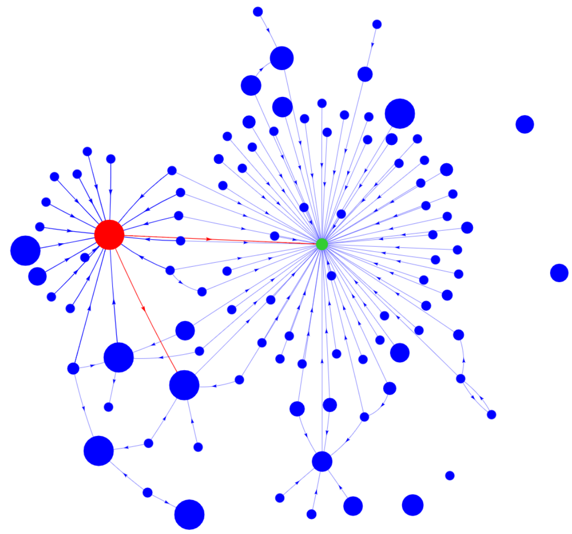
B.1 Force-directed graph drawing

The experimental software uses a force-directed algorithm that imposes both attraction and repulsion forces between nodes in the network, in order to readjust their position in space and improve the overall visibility on the subjects' screen. In this context, the nodes are considered as bodies or particles by the simulation tool described below, and the various forces applied between them can be described as follows:

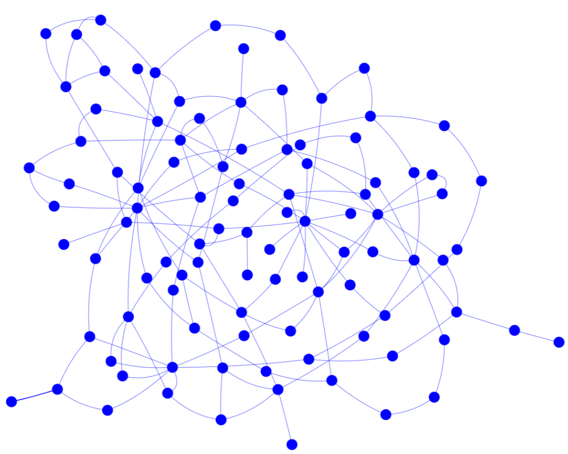
- *Repulsion forces*: all nodes from the network apply a repulsion force F_r to each other to avoid overlaps and allow a sparse visualization of the network. The computation of this force for every node is a complex task, especially in the context of large networks (time complexity is $O(n^2)$ where n is the number of nodes). The software therefore approximate this computation using the well-known Barnes-Hut algorithm introduced by Barnes and Hut [1986], which reduces the time complexity to $O(nlog(n))$.



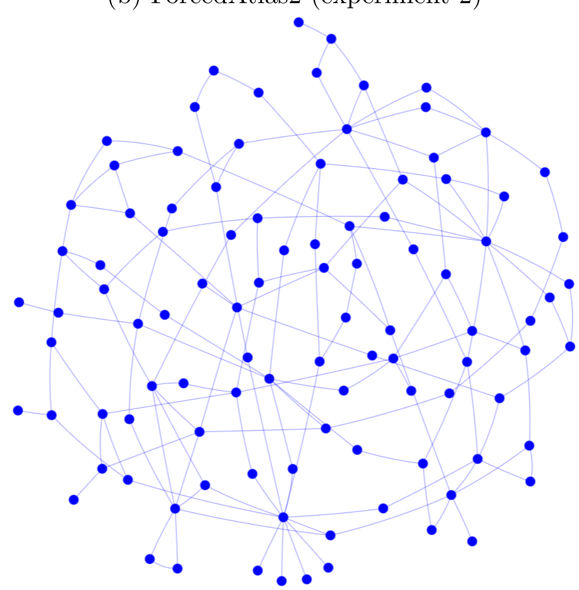
(a) Spring-Electric (experiment 2)



(b) ForcedAtlas2 (experiment 2)



(c) Spring-Electric (experiment 3)



(d) ForcedAtlas2 (experiment 3)

Figure 22: Comparison of network visualization across algorithms (using parameters from Section B.1.4).

- *Attraction forces*: nodes that are linked with each other in the network apply attractive forces F_s towards each other to allow for visual proximity of connected nodes.

- *Gravity force*: all nodes are applied a gravitational force F_g to a center of origin O to pull the entire network towards the center of the screen. In particular, such a force allows disconnected components to be within reasonable distance from each other, and therefore more easily visualized on the screen.

In summary, nodes are attracted by gravity and other nodes they are linked with, and repulsed by other nodes they are not linked with.

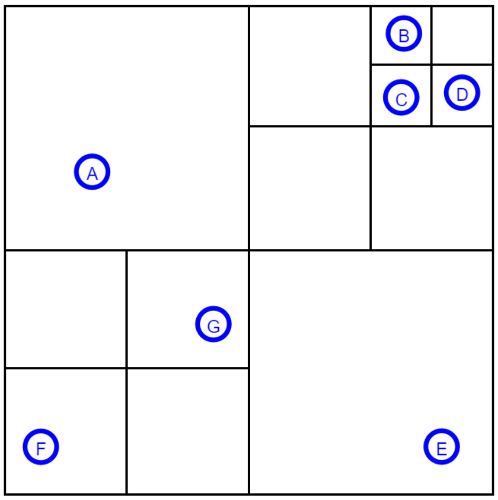
B.1.1 Barnes-Hut approximation algorithm

The Barnes-Hut algorithm used to compute repulsion forces consists in first constructing a quad-tree by recursively dividing the visual space into same size quadrants such that every player node can eventually be associated with exactly one region based on its position in space (leaf of the tree). Figure 23a depicts an example of such recursive division of space, and Figure 23b presents the corresponding quad-tree with all possible quadrants. In this case, nodes A and E are associated with only one iteration of this process, whereas nodes F and G require two such iterations, and nodes B, C, and D required three iterations. Each region of the quad-tree (including at least one node) reached after $p \geq 0$ iterations is associated with two values (s, c):

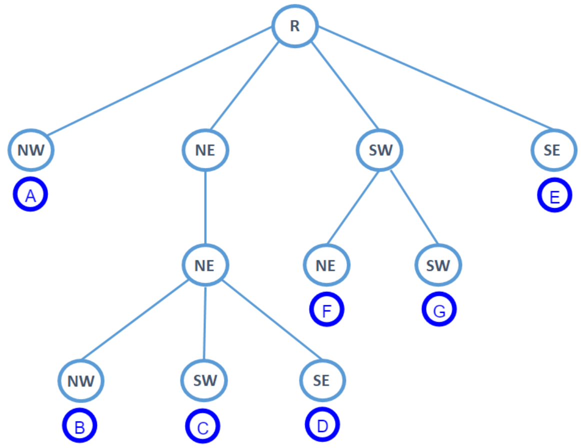
- The size s, corresponding to the width (or height) of the quadrant, i.e., $s = \frac{S}{2^p}$ where S represents the width (or height) of the entire space.

- The center of mass c whose position in the Cartesian coordinate system corresponds to $x_c = \frac{1}{M} \sum_{i=1}^{n} m_i x_i$ and $y_c = \frac{1}{M} \sum_{i=1}^{n} m_i y_i$ where (x_i, y_i) represents the coordinates of node i in space, m_i the mass of node i, and $M = \sum_{i=1}^{n_q} m_i$ the sum of masses among nodes located in the quadrant (n_q). Note that in the presence of a unique node in the region, the center of mass is equivalent to that node's position.

The algorithm then aggregates groups of nodes located in the same region to determine a unique force that approximates the sum of individual forces in that group (as if the group of nodes were a single node). More precisely, starting from the largest region of the Barnes-Hut quad-tree (the root), the algorithm assesses the distance $dist(o, c)$ between a given node o and the center of mass of that region c ($dist(o, c) = \sqrt{(x_o - x_c)^2 + (y_o - y_c)^2}$): if this distance is sufficiently large, then the group of nodes in the corresponding region is considered as a single node, else the process is iterated by considering subregions from the



(a) Space division into quadrants. Players are represented as bodies/particles labeled A-G.



(b) Quad-tree. Root node R represents the entire space; children nodes represent any of the four quadrants: North West (NW), North East (NE), South West (SW), and South East (SE); leaf nodes contain only one body; empty quadrants are ignored.

Figure 23: Illustration of Barnes-Hut algorithm.

tree (nodes sufficiently close to o will therefore be considered independently). Formally, the condition for determining whether node o is sufficiently far from a center of mass c is: $\frac{s}{dist(o,c)} < \theta$ where s is the size associated with the region (see above), and θ is a parameter determining the accuracy of the simulation. A large value of θ increases the speed of the simulation but decreases its accuracy. If $\theta = 0$, no approximation is made by the algorithm and repulsion forces are applied from every individual, which can considerably slow down the simulation but guarantees high accuracy.

B.1.2 Modelling the forces

Using this approximation algorithm, our software calculates the repulsion $F_r(o, c)$ applied to a node o by a group of nodes represented by its center of mass c and its total mass M_c. The default formula (used in experiments 1 and 2), as inspired by the behavior of electrically charged particles, is described as follows:

$$F_r(o,c) = \frac{K_r.M_c.m_o}{dist(o,c)^2} \tag{9}$$

Where K_g captures the gravitational constant such that $K_g < 0$ to obtain the repulsion effect. Note that the default mass of every node in the network (as considered in all our experiments) is 1.

An alternative computation of this repulsion force proposed by Jacomy et al. [2014] (used in experiment 3) is as follows:

$$F_r(o,c) = \frac{K_r.M_c.m_o(d_o+1)}{dist(o,c)} \tag{10}$$

where d_o represents the degree of node o in the network. Note that this definition of the repulsion force slightly differs from Jacomy et al. [2014] in that it only considers the degree of the node of focus o.

Similarly, the attraction force applied between two linked nodes o_1 by o_2 follows the attraction formula of springs, which is defined as follows:

$$F_s(o_1,o_2) = K_s.(dist(o_1,o_2) - L) \tag{11}$$

Where L defines the resting length of an edge, and K_s the spring gravity constant such that $K_s > 0$ to obtain the attraction effect. Note from Equation (11) that the force applied on two linked nodes is symmetric, i.e., both nodes are equally attracted to each other.

Finally, the central gravity force applied to node o is computed as follows (default formula used in experiments 1 and 2):

$$F_g(o) = K_g.dist(o,O).m_o \tag{12}$$

Where O represents the position of the point of origin, and K_g the central gravity constant such that $K_g > 0$ to obtain the attraction effect.

An alternative computation of this central gravity force relying on the method by Jacomy et al. [2014] (used in experiment 3) is as follows:

$$F_g(o) = K_g.(d_o + 1).dist(o, O).m_o \tag{13}$$

The net force vector applied to any node o resulting from the above three forces is then:

$$F_x(o) = \frac{x_O - x_o}{dist(o, O)} F_g(o) + \sum_{o' \in N \setminus \{o\}} \frac{x_{o'} - x_o}{dist(o, o')} F_a(o, o') + \sum_{c \in C_o} \frac{x_c - x_o}{dist(o, c)} F_r(o, c) \tag{14}$$

$$F_y(o) = \frac{y_O - y_o}{dist(o, O)} F_g(o) + \sum_{o' \in N \setminus \{o\}} \frac{y_{o'} - y_o}{dist(o, o')} F_a(o, o') + \sum_{c \in C_o} \frac{y_c - y_o}{dist(o, c)} F_r(o, c) \tag{15}$$

Where C_o represents the set of centers of mass associated with regions in the Barnes-Hut quad-tree where the condition $\frac{s}{dist(o,c)} < \theta$ (see details in the previous section).

B.1.3 Dynamics

The above static properties describe the net forces that are applied in the network, given the positions of all nodes and the links between nodes. The resulting dynamic update of the network is achieved by computing the corresponding velocity of nodes on both coordinate axes. More precisely, the velocity applied to a node o at a time t on both coordinate axes (x and y) is determined as follows:

$$V_x(o, t) = max(V_{max}, \frac{F_x(o) - D.V_x(o, t-1)}{m_o}.T + V_x(o, t-1)) \tag{16}$$

$$V_y(o, t) = max(V_{max}, \frac{F_y(o) - D.V_y(o, t-1)}{m_o}.T + V_y(o, t-1)) \tag{17}$$

Where D represents the damping factor determining how much of the velocity from the previous simulation iteration carries over to the next iteration, T the time step for the discrete simulation, and V_{max} the maximum velocity of nodes (used to increase time to stabilization). We assume no initial velocity, i.e., $V_x(o, 0) = V_y(o, 0) = 0$. Given such computed velocity, the position update of a node o at any time t directly follows:

$$x_o^t = x_o^{t-1} + V_x(o, t).T \tag{18}$$

$$y_o^t = y_o^{t-1} + V_y(o, t).T \tag{19}$$

The discrete simulation terminates and node o stabilizes whenever the associated velocity becomes sufficiently low with respect to some given threshold (V_{min}). More precisely, the convergence rule for every node o is:

$$\sqrt{V_x(o,t)^2 + V_y(o,t)^2} < V_{min} \tag{20}$$

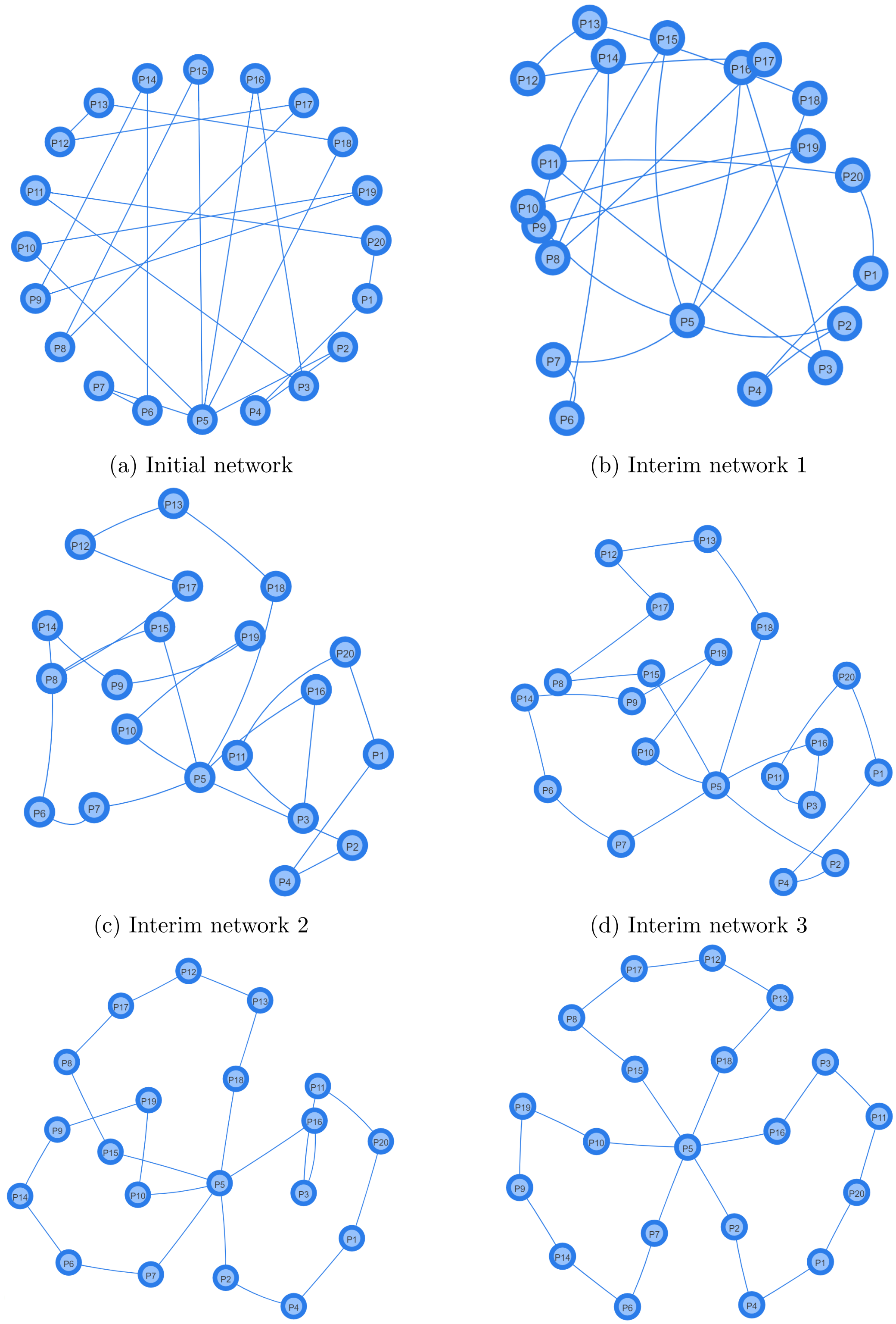
The following figure, as an illustration of the dynamics, shows the dynamic updates of the network in Figure 1, moving from the fixed visualization (Figure 24a) to the adaptive visualization (Figure 24f).

B.1.4 Model parameters

Model parameter settings used across experiments are described in Table 2.

	Experiments		
	Linking Game	Connectors & Influencers	Brokerage & Market Power
θ	0.5	0.5	0.5
K_g	-2000	-2000	-50
K_s	0.04	0.04	0.08
K_g	0.3	0.3	0.01
L	95	95	50
D	0.09	0.09	0.8
T	0.5	0.5	0.5
V_{min}	0.3	0.3	0.3
V_{max}	10	10	50
m_i	1	1	1

Table 2: Parameter values in experiments



(a) Initial network

(b) Interim network 1

(c) Interim network 2

(d) Interim network 3

(e) Interim network 4

(f) Final network

Figure 24: Dynamic adjustment of networks using the force-directed algorithm

C Experimental Instructions and Tutorials

C.1 Linking Game

[In the following instructions, N is to be replaced with a value from $\{10, 50, 100\}$*, R with a value from* $\{40, 200, 400\}$*, C with a value from* $\{20, 100, 200\}$*, and E with a value from* $\{50, 250, 500\}$*, depending on the treatment]*

Please read the following instructions carefully. **These instructions are the same for all the participants**. The instructions state everything you need to know in order to participate in the experiment. If you have any questions, please raise your hand. An experimenter will answer your question.

You can earn money by earning points during the experiment. The number of points that you earn depends on your own choices and the choices of other participants. At the end of the experiment, the total number of points that you have earned will be exchanged at the following rate:

$$R \textbf{ points} = \textbf{1 Euro}$$

The money you earn will be paid out in cash at the end of the experiment. The other participants will not see how much you earned.

Details of the experiment

The experiment consists of 6 (six) independent rounds of the same form. The first round is for practice and does not count for your payment. The next 5 rounds will be counted for your payment. At the beginning of each round, you will be grouped with $N-1$ other participants. This group will remain fixed throughout the 6 rounds. Each of the other participants will be randomly assigned an identification number of the form "Px" where x is a number between 1 and $N-1$. Those numbers will be randomly changed across every round of the experiment. The actual identity of the participants will not be revealed to you during or after the experiment. The participants will always be represented as blue circles on the decision screen. You are always represented as a yellow circle identified as "ME".

Each round will last for **6 (six) mins: the first minute will be a trial period, only the subsequent 5 minutes will be relevant for the earnings**. Your earnings in a given round will be based on everyone's choice **at a randomly selected moment in the last 5 minutes of the round**. In other words, any decision made before or after that randomly chosen moment will not be used to determine points that you earn in the round. This precise moment will be announced to everyone only at the end of the round, along with your choice and that of others connected with you in the network at that moment.

At the beginning of the experiment, you are given **an initial balance of E points**. Your final earnings at the end of the experiment will consist of the sum of points you earn across the 5 last rounds plus this initial balance (the first round will be used to familiarize yourself with the game and will have no influence on your earnings). Note that if your final earnings (i.e., the sum of your earnings across the 5 last rounds plus the initial balance) go below 0, your final earnings will be simply treated as 0.

In each round, every participant will be allowed to form links with other participants or delete links that were previously created by him- or herself at any moment during the 6 minutes. You are linked with another person if you form a link with that person or that person forms a link with you (or both). Each link you form costs you C **points**. You do not pay any cost for links formed by others. In order to form or delete a link with a participant, you will simply need to double-click the corresponding node on the computer screen. A network resulted from your choice and choices of other participants at any moment will be updated in your computer screen in real time.

The participants that you are linked with (regardless of whether you or they form the links) are called your neighbours. You are said to be **connected** with another participant when there exists a sequence of links connecting you with that person in the network.

The computer screen will be split into two parts:

- **The left side of the screen presents you and participants that you are connected with**.
- **The left side of the screen presents you and participants that you are connected with**.

Each node is described by their identification number "Px". Identification numbers "Px" are randomly assigned in every round. Therefore, every player is likely to have a different ID in different rounds.

At the very beginning of the round when no link is formed, you will be on the left side of the screen and all the other participants will be shown on the right side of the screen.

You may revise your choices at any moment before the round ends. During a round, you will also be informed about every other participant's most recent decision (formed links), which will be updated every 2 seconds or whenever you change your own choice.

Earnings

Your earnings at any moment of the round are determined by **the benefits that you obtain** *minus* **the costs that you incur** from the network at that moment.

The costs that you incur from the network are equal to C points times the number of links created by you.

The benefits that you obtain from the network are equal to the sum of benefits you receive from each of the other participants to whom you are connected, plus 10 points.

The benefit you receive from each participant depends on the distance between you and that participant. This distance is defined by the smallest number of links that connect you with the participant in the network. For example, the distance between you and each of your neighbours is 1. The distance between you and each neighbour of your neighbours (who is not your neighbour) is 2.

You receive a benefit of 10×0.9^d points from a participant who is connected with you from a distance d. Given this form of benefit, for example, you receive a benefit of $10 \times 0.9^1 = 9$ points for each of your neighbours, and $0 \times 0.9^2 = 8.1$ points for each neighbour of your neighbours that is not your neighbour. You receive 0 point from each participant whom you are not connected with.

One moment in the last 5 minutes of the round will be randomly chosen to determine every participant's real earnings in the round.

Tutorial

Please follow this simple tutorial simulating a simple virtual scenario on the computer screen. In this tutorial you are interacting with 9 other players, and every link you form costs you 20 points. **Note** that this setting is only illustrative and slightly differs from the real game described above (you will then interact with $N - 1$ other players and every link will cost you C points). In the initial state, you are not linked with anyone: you start with 10 points.

- Initially, the nodes on the right side of the screen represent all other players (in this simulation, those players are not real people). You may choose to form a link with any player by simply double clicking on the corresponding node. For example, forming a link with P4 reveals that each of P2 and P3 forms one link with P4. Forming a link with P4 costs you 20 points (in red on the screen), but it also generates a benefit of 35.2 points (10 + 9 (= 10×0.9^1 from P4) + 8.1 (= 10×0.9^2 from P2) + 8.1 (= 10×0.9^2 from P3). Your resulting earnings are **15.2 points (= 35.2 points - 1 link 20 points)**.

- After forming a link with P4, you observe that some nodes remain not connected with you (P1, P5, P6, P7, P8, and P9 on the right side). However, forming an additional link with P9 (by double clicking on the corresponding node) reveals that all those nodes were connected with one another and that you are now connected with every participant. You were not allowed to observe them before because they were not linked with any node you were connected to. You can now observe them because there exists a sequence of links connecting you from any of them (for example, P5 is connected to you via P9). Remember that you can only see players that are connected with you. Your resulting earnings become **44.7 points (= 84.7 points - 2 links 20 points)**.
- Alternatively, you may choose to remove a link that you previously formed by double clicking on the corresponding node. For example, after forming links with P4 and P9, removing the link with P4 makes players P2, P3, and P4 move to the right side of the screen, as they are not connected with you anymore.
- You may also shape the visual structure of the network by dragging nodes as it pleases you.

Summary

Here is a brief description of information available on the decision screen:

1. The timer indicates elapsed time since the beginning of the round. Any round lasts **6 minutes**. A moment will be randomly selected **in the last 5 minutes** to determine everyone's payoff. The time displayed will turn red when entering this interval.
2. **Only decisions made at the randomly selected moment in the round** matter to directly determine the earnings. The payoff may be negative at the end of a round. However, starting from a balance of E points, any negative total of points at the end of the 5 rounds will be equivalent to 0 point.

3. A participant is connected with you if there exists a sequence of links connecting you to that person in the network.

4. For every participant you are connected with, you receive 10×0.9^d points where d represents the smallest number of links that connect you with that person in the network.

5. However, you receive 0 points from every participant you are not connected with. For every link you form, you pay C points.

6. You are represented as the yellow node, and your ID is “ME”.

7. Every other node’s ID is represented as “Px” (inside the node) where x is a number. Every node has a unique ID, which is randomly reassigned in every round.

C.2 Connectors and Influencers

[In the following instructions, N is to be replaced with any value from $\{3, 7, 49, 99\}$ *depending on the treatment]*

Please read the following instructions carefully. **These instructions are the same for all the participants**. The instructions state everything you need to know in order to participate in the experiment. If you have any questions, please raise your hand. One of the experimenters will answer your question.

You can earn money by earning points during the experiment. The number of points that you earn depends on your own choices and the choices of other participants. At the end of the experiment, the total number of points that you have earned will be exchanged at the following exchange rate:

100 points = 1 Euro

The money you earn will be paid out in cash at the end of the experiment. The other participants will not see how much you earned.

Details of the experiment

The experiment consists of 6 (six) independent rounds of the same form. The first round is for practice and does not count for your payment. The next 5 rounds will be counted for your payment.

At the beginning of each round, you will be grouped with N other participants. This group will remain fixed throughout the 6 rounds. Each of the participants will be randomly assigned an identification number of the form "Px" where x is a number between 1 and N. Those numbers will be randomly changed across every round of the experiment. The actual identity of the participants will not be revealed to you during or after the experiment. The participants will always be represented as blue circles on the decision screen. You are always represented as a yellow circle identified as "ME".

Each round will last **6 (six) mins: the first minute will be a trial period, only the latter 5 minutes will be relevant for the earnings**. Your earnings in a given round will be based on everyone's choice **at a randomly selected moment in the last 5 mins of the round**. In other words, any decision made before or after that randomly chosen moment will not be used to determine your points. This precise moment will be announced to everyone only at the end of the round, along with the corresponding behavior and earnings.

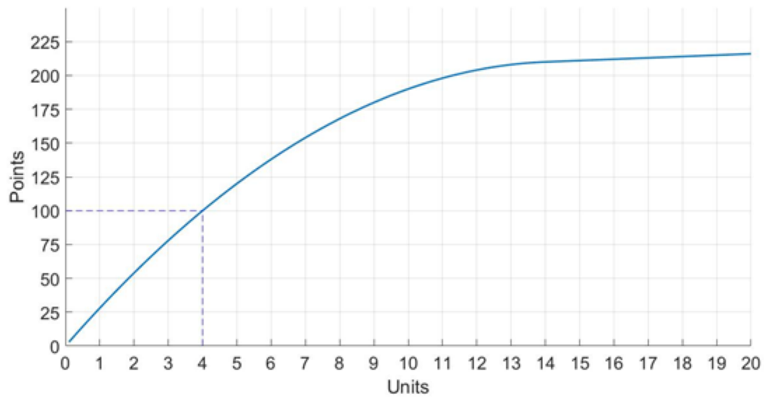
At the beginning of the experiment, you are given an initial balance of 500 points. Your final earnings at the end of the experiment will consist of the sum of points you earn across the 5 last rounds plus this initial capital (the first round will be used to familiarize yourself with the game and will have no influence on your earnings). Note that if your final earnings (i.e., the sum of your earnings across the 5 last rounds plus the initial endowment) go below 0, your final earnings will be simply treated as 0.

In each round, every participant will have choose two types of actions:

- **How many any units to buy/invest**: You may buy at most 20 units. Each unit costs you **11 points**.
- **Add/delete links with other participants**: You are linked with another person if you form a link with that person or that person forms a link with you (or both). You do not pay any fee for links formed by others. The people that you are linked with (regardless of whether you or they form the links) are called your neighbours. You automatically have access to **all units bought by your neighbours** as well as **half of the units bought by your neighbours' neighbours** (see below for an example). Each link you form costs you **95 points**.

You may revise your choices at any moment before the round ends. During a round, you will also be informed about every other participant's most recent decision (units bought and formed links), which will be updated every 5 seconds or whenever you change your own choice.

At any moment, the total number of units you have access to (i.e., units you bought + units bought by your neighbours + units bought by your neighbours' neighbours) generates points for you according to the following figure (for example, accessing 4 units generates 100 points, as shown by the dotted lines):



Moreover, having access to 20+m units generates 216+m points.

The computer screen will be split into two parts:

- **The middle side of the screen presents you and your local neighbourhood**. More precisely, you will see your neighbours, the neighbours of your neighbours, and the neighbours of neighbours' neighbours. In other words, you will see the participants that are up to 3 links away from you.

- **The right side of the screen presents participants outside of your local neighbourhood**.

- **The left side of the screen presents the code for the players' net earnings in the network**. *[Payoff information treatment only]* The inner circle of each node from the middle or right part side of the screen is characterized by some color, which varies from **green** (high positive net payoff) to **red** (high negative net payoff) depending on the player's corresponding net earnings.

Each node is described by their identification number "Px" and the number of units that they buy. Identification numbers "Px" are randomly assigned in every round. Therefore, every player is likely to have a different ID in different rounds. In the initial state of the network, nobody buys any unit and no link is formed.

Tutorial

Please follow this simple tutorial simulating a simple virtual scenario on the computer screen. In this tutorial you are interacting with 9 other players. In the initial state, you are not linked with anyone and you do not buy any units: you start at 0 points.

1. The slider allows you to choose how many units you wish to buy yourself. For example, buying 4 units costs you 44 points (= 4 units × 11 points, in red on the screen) and generates 100 points (according to the figure from the previous page, in green on the screen).

2. Initially, the nodes on the right side of the screen represent all other players (in this simulation, those players are not real people). The size of node reflects the total number of units bought by that node and the units accessed via the network. For example, P1-P4 are the largest nodes because these players have access to the most units.

3. You may choose to form a link with any player by simply double clicking on the corresponding node. For example, forming a link with P4 reveals that P1, P2, and P3 each form a link with P4, and P9 forms a link with P1. Forming a link with P4 costs you 95 points (in red on the screen), but it also gives you access to 8.5 units (7 from P4 + 0.5 × 1 from P1 + 0.5 × 1 from P2 + 0.5 × 1 from P3), which generates 174 points (according to the above figure, describing the benefit function in green on the screen). If you do not buy any additional unit yourself, your resulting net payoff is **79 points (= 174 points − 1 link × 95 points)**.

4. After forming a link with P4, you observe that some nodes remain unobserved (P5, P6, P7, and P8 on the right side). However, forming an additional link with P9 (by double clicking on the corresponding node) reveals that those nodes all form a link with P9. You were not allowed to observe them before because they were 4 nodes

away from you (for example, P5 were connected to you via P4, P1, and P9). You can now observe them because they are only 2 nodes away from you (for example, P5 is connected to you via P9 only). Remember that you can only see players that are at most 3 nodes away. Assuming you still do not buy any unit yourself, your resulting net payoff is **16 points (= 206 points from accessing 12.5 units − 2 links × 95 points)**.

5. Alternatively, you may choose to remove a link that you previously formed by double clicking on the corresponding node. For example, after forming links with P4 and P9, removing the link with P4 leads to players P2 and P3 becoming unobserved again, as they are now more than 3 nodes away from you.

6. Note that varying the amount of units you buy directly affects the sizes of the nodes you are linked with as well as their neighbours. Indeed, the amount of units they each have access to includes the units you buy (the larger this amount, the larger the node).

7. You may also shape the visual structure of the network by dragging nodes as it pleases you.

Summary

Here is a brief description of information available on the decision screen:

1. The timer indicates elapsed time since the beginning of the round. Any round lasts **6 mins**. A moment will be randomly selected **in the last 5 mins** to determine everyone's payoff. The time displayed will turn red when entering this interval.

2. **Only decisions made at the randomly selected moment in the round** matter to directly determine the earnings. The payoff may be negative at the end of a round. However, starting from a balance of 500 pts, any negative total of points at the end of the 5 rounds will be equivalent to 0 point.

3. The amount of units you have access is equal to the sum of **(1)** the units bought by you, **(2)** the units bought by your neighbours, and **(3)** half of the units bought by your neighbours' neighbours.

4. You are represented as the yellow node, and your ID is "ME".

5. Every other node's ID is represented as "Px" (inside the node) where x is a number. Every node has a unique ID, which is randomly reassigned in every round.

6. The size of each node determines **how many units that node has access to** (units bought personally plus units accessed from others, directly and indirectly).

7. The amount of units **bought personally by** a player is mentioned inside the corresponding node.

8. *[Payoff information treatment only]* The color of each node determines **that node's net earnings** according to the code depicted on the left side of the screen.

C.3 Brokerage and Market Power

[In the following instructions, N is to be replaced with a value from $\{10, 50, 100\}$, R with a value from $\{20, 110, 220\}$, C with a value from $\{8, 40, 80\}$, and E with a value from $\{40, 200, 400\}$, depending on the treatment]

[All treatments]
Please read the following instructions carefully. **These instructions are the same for all the participants**. The instructions state everything you need to know in order to participate in the experiment. If you have any questions, please raise your hand. One of the experimenters will answer your question.

In addition to the 5 euro show up fee that you are guaranteed to receive, you can earn money by scoring points during the experiment. The number of points depends on your own choices and the choices of other participants. At the end of the experiment, the total number of points that you have earned will be exchanged at the following exchange rate:

$$R \textbf{ points} = \textbf{1 Euro}$$

The money you earn will be paid out in cash at the end of the experiment. The other participants will not see how much you earned.

In this experiment, you will participate in 6 independent rounds of the same form. The first round is for practice and does not count for your payment. The next 5 rounds will be counted for your payment. At the beginning of the first round, you will be grouped with

$N-1$ other participants; so there are N participants in all in your group. This group will remain fixed throughout the six rounds.

A round

We now describe in detail the process that will be repeated in each of the six rounds.

At the beginning of a round, you in your computer screen will be identified as the circle of 'Me' and the other participants will be randomly assigned an identification number of the form "Px" where x is a number between 1 and $N-1$, and identified as the circle of "Px". The ID assignment of the other participants will remain unchanged within the round and will be randomly made again at the beginning of the next round (e.g., node P4 does not refer to the same participant across different rounds).

Each round will last 6 (six) minutes. At the very beginning of the round, participants will start with an empty network where no link among them is formed. All participants will then be asked to propose any number of links to any of the other participants to whom they wish to link by double-clicking on their corresponding nodes. Anyone who makes a link proposal to you (while you do not make a link proposal with them) will become **triangle-shaped**. For example, players P2, P3, and P4 make link proposals to you in the left part of Figure 1 (while you do not make any link proposal). Similarly, any link proposal that you make to player who does not make one with you will become **square-shaped**.

A link between two participants will be formed only if both of the participants proposed a link with each other. Anyone who is linked with you, called your neighbour, will become **circle-shaped**.

For example, from the left part of Figure 1, suppose that you make link proposals to each of P2, P3, P4 and P6. Because you also received link proposals from P2, P3, and P4, each of them is now linked to you and becomes circle-shaped. This is shown in the right part of Figure 1. On the other hand, P6 did not make a link proposal to you, and as a result, P6 will become square-shaped on your screen. Those who neither proposed a link to you nor received a link proposal from you will remain circle-shaped (for example, P1, P5, and P7, in the right part of Figure 1). Note that the network depicted in Figure 1 is also shown on your screen as a tutorial for you to test the experimental interface by creating and/or removing link proposals with other (virtual) players.

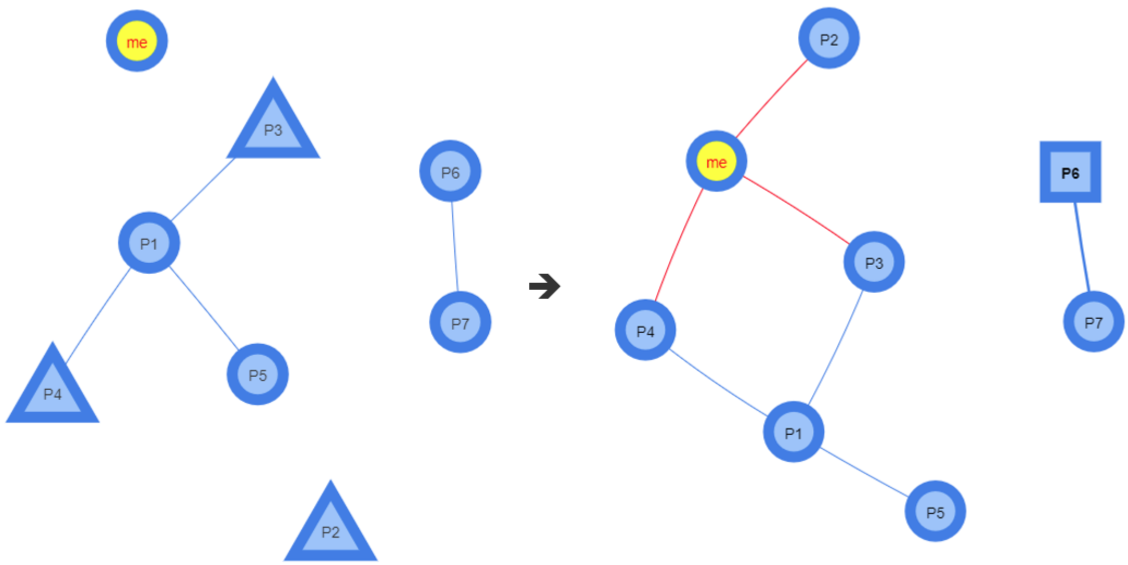


Figure 1

Every participant will be allowed to **add/delete link proposals with other participants** at any moment during the six minutes of the round. If you delete a link proposal to a participant who was linked to you (someone shown circle-shaped in the screen), then that participant will no longer be linked to you and will revert to being triangle-shaped. If you delete a link proposal to a participant who received a link proposal from you but did not propose a link to you (who was shown square-shaped in the screen), that participant will become circle-shaped. For example, starting from the network in the left part of Figure 2, deleting the link proposals to P3 and P6 will result in the network shown on the right part of Figure 2. In these ways, the computer screen will update the network every 2 seconds or whenever you revise your linking decision.

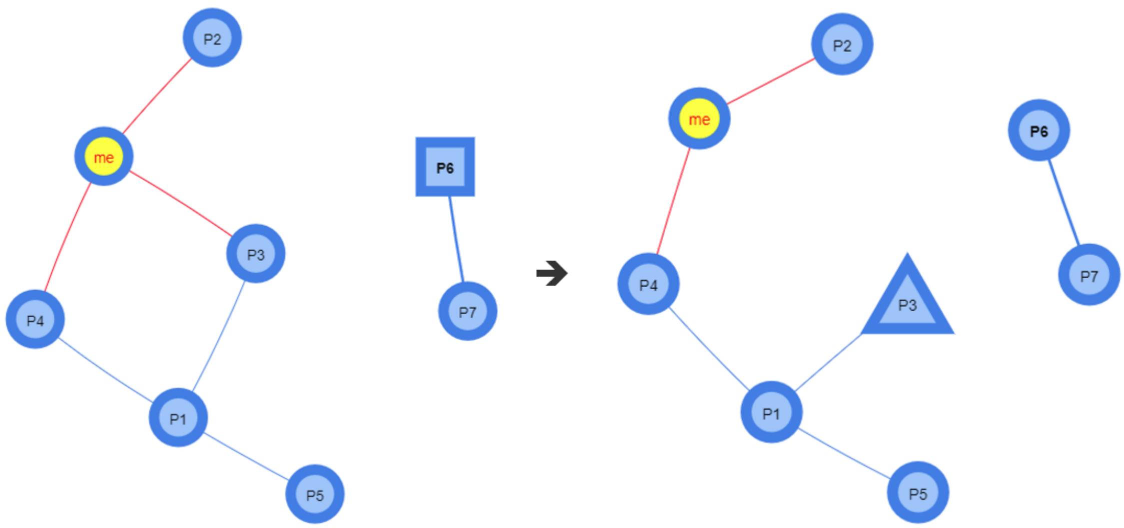


Figure 2

In summary, the shape of each participant on the computer screen indicates your relationship with them.

- **Circle**: they are linked with you or unlinked (no proposal from them or from you)
- **Square**: you propose a link, but they do not reciprocate.
- **Triangle**: they propose a link to you, but you do not reciprocate.

The first minute of each round will be a trial period and only the last 5 minutes will be relevant for your earnings in that round. Your earnings in the round will be based on everyone's choice **at a randomly selected moment in the last 5 minutes of the round**. In other words, any decision made before or after that randomly chosen moment will not be used to determine your points. This precise moment will be announced to everyone only at the end of the round, along with the corresponding behavior and earnings.

In order to help you keep track of potential earnings which you and the other participants make during the round, your earnings at each moment will be presented at the top part of the computer screen. In addition, the payoff of each participant from the network is presented inside their corresponding node (rounded to the closest integer, below their identification number).

After participants are informed of their earnings at a randomly chosen moment, the next round will start with the computer randomly assigning IDs of the other participants in your group. This group is the same as in the previous round. However, IDs of other players are likely to be different across different rounds.

Earnings

At the beginning of the experiment, you are given **an initial balance of E points**. The first round will be used to familiarize yourself with the experiment and will have no influence on your earnings. Your final earnings at the end of the experiment will consist of the sum of points you earn across the last 5 rounds, plus this initial balance. Note that if your final earnings go below 0, they will be treated as 0.

Your earnings in each round depend on benefits you get from your own connection to the other participants and whether you are critical for the connection between two other participants (brokerage rent), and the cost of linking you pay.

In a network, two participants are said to be **connected** when there exists a path linking them. For example, in the right part of Figure 1, you are connected with the five participants of P1, P2, P3, P4, and P5.

[Treatment Criticality only]
A participant is said to be **critical** for the connection between two other participants if they are connected and the participant lies on ALL paths between them. In the pair between you and P5, P1 is critical because P1 lies on each of the two paths between you and P5.

Every connected pair of two participants creates a value of **10 points**. The pair creating the value of 10 points shares this value equally among themselves and all the critical participants between them.

Your total benefits consist of

(1) The benefits you earn from your own connection to other participants,

(2) The brokerage rents you earn from by being critical for the connection between pairs of other participants.

[Treatment Betweenness only]

Flow of transactions between two participants will be only made through a shortest path between them in the network. This means that only a participant who lies on a shortest path between two other participants can be involved in transactions and earn brokerage rent. For instance, consider the right part of Figure 1: between Me and P5, there are two shortest paths: Me-P4-P1-P5 and Me-P3-P1-P5. Both paths have two participants lying on them, and can be used for trade between Me and P5.

Every connected pair of two participants create a value of **10 points**. This value is divided equally among the connected pair and participants lying on any existing shortest path. If M is the number of participants lying on any shortest path for the pair, then each member of the pair earns **10/(M+2) points**. Other participants lying on any shortest path earns **10/(M+2) points** multiplied by the proportion of the number of shortest paths that she lies on. By way of illustration, consider the right part of Figure 1: there are two shortest paths between Me and P5 with 2 participants lying on each of them (M=2), and therefore participants P3 and P4 who lie on one shortest path each receive 1.25 points $(10/(2+2) \times 1/2)$. However, participant P1 lies on both the shortest paths and receives 2.5 points $(10/(2+2))$.

Your benefits therefore consist of

(1) The benefits you earn from your own connection to other participants,

(2) The brokerage rents you earn for lying on shortest paths between pairs of other participants.

[All treatments]

On the cost side, you pay C **points** per link that is created by you. Note that a link proposal made by you will cost you C points only if the participant who received your link proposal has also made a link proposal to you. Otherwise, your link proposal does not create a link and costs nothing.

Therefore, your earnings in each round correspond to the network chosen at a random moment from the last five minutes of the experiment.

Earnings = (sum of values you obtain from your connections with others) + (sum of values you obtain from brokerage) – (total cost of links created by you)

The top part of the computer screen shows you the earnings that are decomposed into the three parts:

- **Benefits from being connected**
- **Brokerage rents**
- **Costs of linking**

To give you a concrete idea of how each part of earnings is computed, let us take the network on the right part of Figure 1. You are presented as Me.

First, observe that there is no path to two participants - P6 and P7: as you are not connected to them, you obtain no benefit from them.

[Treatment Criticality only]
Second, you are connected to four participants – P1, P2, P3 and P4 – without any critical participants. You obtain $10/2 = 5$ points from each of these connections. You are also connected to another participant, P5, and there is one critical participant, P4, between you and P5. You obtain $10/3 = 3.3$ points. Therefore, the benefits that you get from your connections are

$$\frac{10}{2} \times 4 + \frac{10}{3} \approx 23.3$$

Third, observe that you are critical between P2 and each of the four participants – P1, P3, P4, and P5. So you obtain brokerage rents from these pairs. Specifically, you are the only critical participant in three pairs – (P2, P1), (P2, P3), and (P2, P4). In the pair (P2, P5), you and P1 are both critical. The brokerage rents you obtain are

$$\frac{10}{3} \times 3 + \frac{10}{4} = 12.5$$

Recall that your payoff is only affected by the reciprocated links.

[Treatment Betweenness only]
Second, you are connected to three participants – P2, P3 and P4 – without any intermediary. You obtain $10/2 = 5$ points from each of these connections.

You are connected to participant P1, through two participants P3 and P4, lying on two distinct shortest paths, between you and P1. You and participant P1 each receive $10/(2+1) = 3.3$ points. You are also connected to participant P5 through 3 intermediaries: P3 and P4 lying on only one shortest path, and P1 lying on both shortest paths. You and participant P5 each receive $10/(2+2) = 2.5$ points. Therefore, the benefits that you get from your connections are

$$5 \times 3 + 3.3 + 2.5 \approx 20.8$$

Third, observe that you lie on all shortest paths between P2 and each of the four participants – P1, P3, P4, and P5. So you receive brokerage rents from these pairs – (P2, P1), (P2, P3), (P2, P4), and (P2, P5). You are the only intermediary for the pairs (P2,P3) and (P2,P4) and therefore earns 3.3 points ($\approx 10/(1+2)$) for each.

Two other intermediaries (P3 and P4) are lying on a shortest path for the pair (P2, P1). Since they each lie on only one of the two existing shortest paths, you earn 2.5 points ($= 10/(2+2)$).

Similarly, there are three other intermediaries lying on a shortest path for the pair (P2, P5): P3 and P4 lie on only one of the two shortest paths, and P1 lies on both of them (as you do). As a result, you earn 2 points ($= 10/(3+2)$).

You also lie on one shortest path out of the two shortest paths between P3 and P4. You therefore receive brokerage rents 1.7 points ($\approx 10/(1+2) \times 1/2$) from this pair. The total brokerage rents you obtain are

$$2 \times 3.3 + 2.5 + 2 + 1.7 \approx 12.8$$

Recall that your payoff is only affected by the reciprocated links.

D Network game interface

D.1 Linking Game

The decision making interface used in the experiment is similar across all treatments. More specifically, Figure 25 illustrates a (fictitious) example of a subject's computer screen in the treatment with $N = 100$. The top part of the screen depicts information about the timer indicating how much time has lapsed in the current round (the timer turns red when payoffs become effective, i.e., after more than 1 minute), and a comprehensive description of the subject's own payoff. Information about payoffs include gross earnings (from connections with others), the cost of linking (number of links formed multiplied by k), and the net earnings (costs substracted from gross earnings). The bottom part of the screen shows detailed information about the network (the subject's node is highlighted in yellow): the subject's local network is represented on the left (entire connected component), other players outside of the subject's local network are found on the right. Note that a scrolldown feature is available for the subject to explore every player outside of his/her local network. Baseline treatments with smaller group sizes use the very same interface (the scrolldown feature is not available then because all players are then directly visible on the screen).

D.2 Connectors and Influencers

The decision making interface used in the experiment is similar across all treatments. More specifically, Figure 26 illustrates a (fictitious) example of a subject's computer screen in the Baseline treatment with $N = 100$. The structure of the screen is similar to the Linking Game experiment described in the previous section. The top part of the screen further includes the subject's own effort, which can be modified via the slider. Information about payoffs here include gross earnings (output of function $f(.)$), the cost of effort (own effort multiplied by c), the cost of linking (number of links formed multiplied by k), and the net

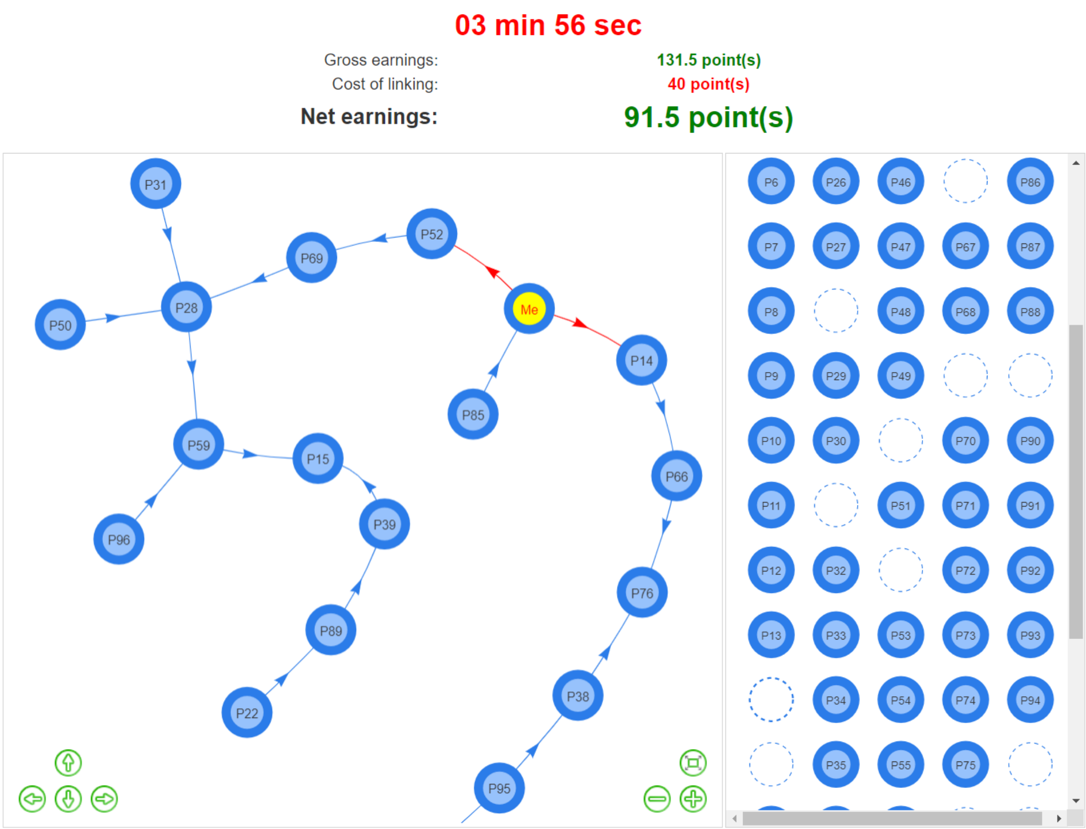


Figure 25: Example of decision screen: Linking Game experiment

earnings (costs substracted from gross earnings). The subject's local network depicted on the bottom left part is restricted to players within a distance 3.

Similarly, Figure 27 illustrates a (fictitious) example of a subject's computer screen in the PayInfo treatment with $N = 100$. The only difference with the decision screen from Figure 26 is about the wider range of colors used to represent the border of each node depicted in the network. Any given node's color is directly associated with that node's corresponding payoff, according to the scale presented on the left part of the screen. payoff-information treatments with smaller group sizes use the very same interface.

D.3 Brokerage and Market Power

The decision making interface used in the experiment is similar across all treatments. More specifically, Figure 28 illustrates a (fictitious) example of a subject's computer screen in Treatment **Criticality** with $N = 100$. Information about payoffs on the top part of the screen include own benefits from own connections, brokerage rents, the cost of linking, and the net earnings. The bottom part of the screen shows detailed information about the entire network (the subject's node is highlighted in yellow).

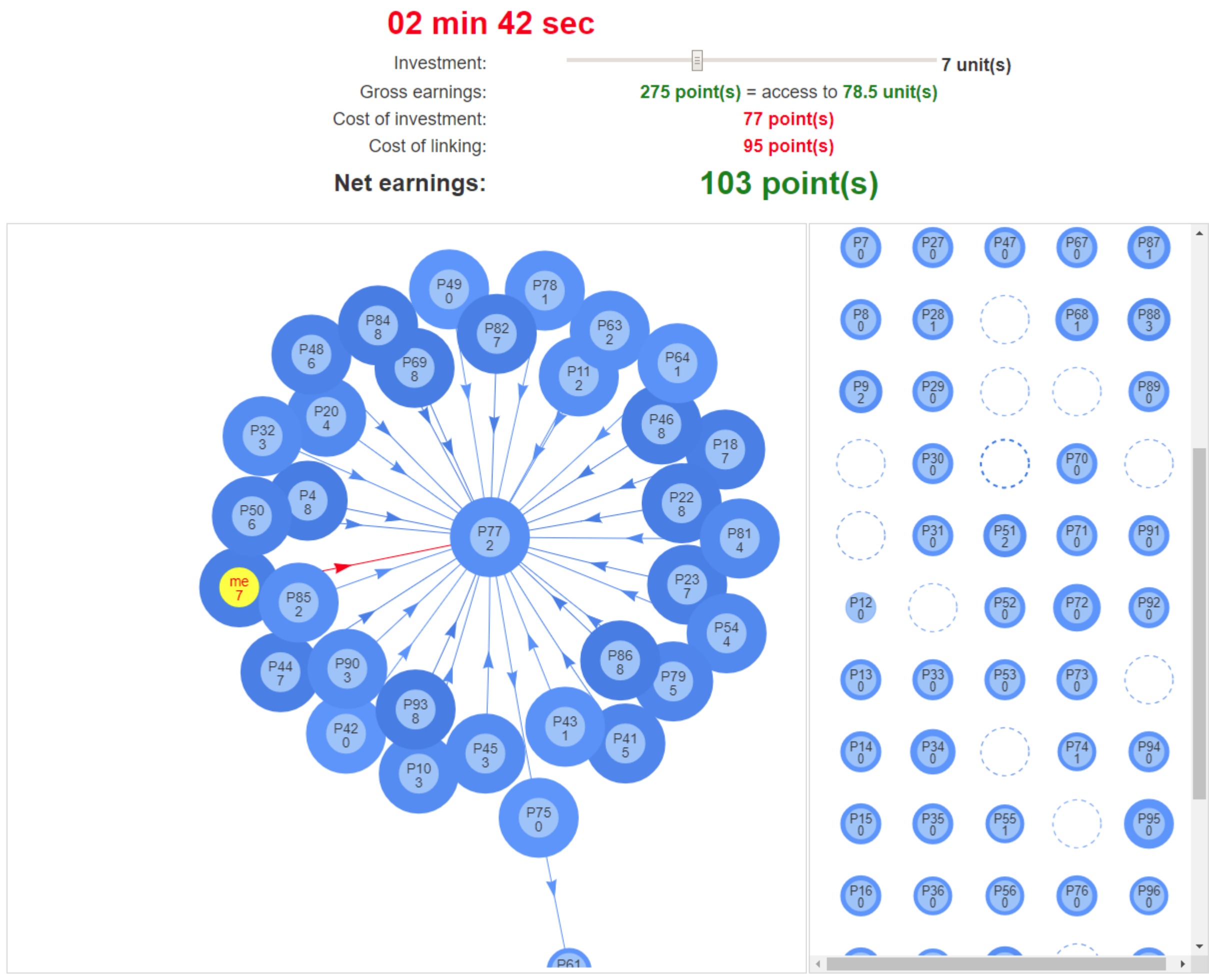


Figure 26: Example of decision screen: Connectors and Influencers experiment (Baseline100 treatment)

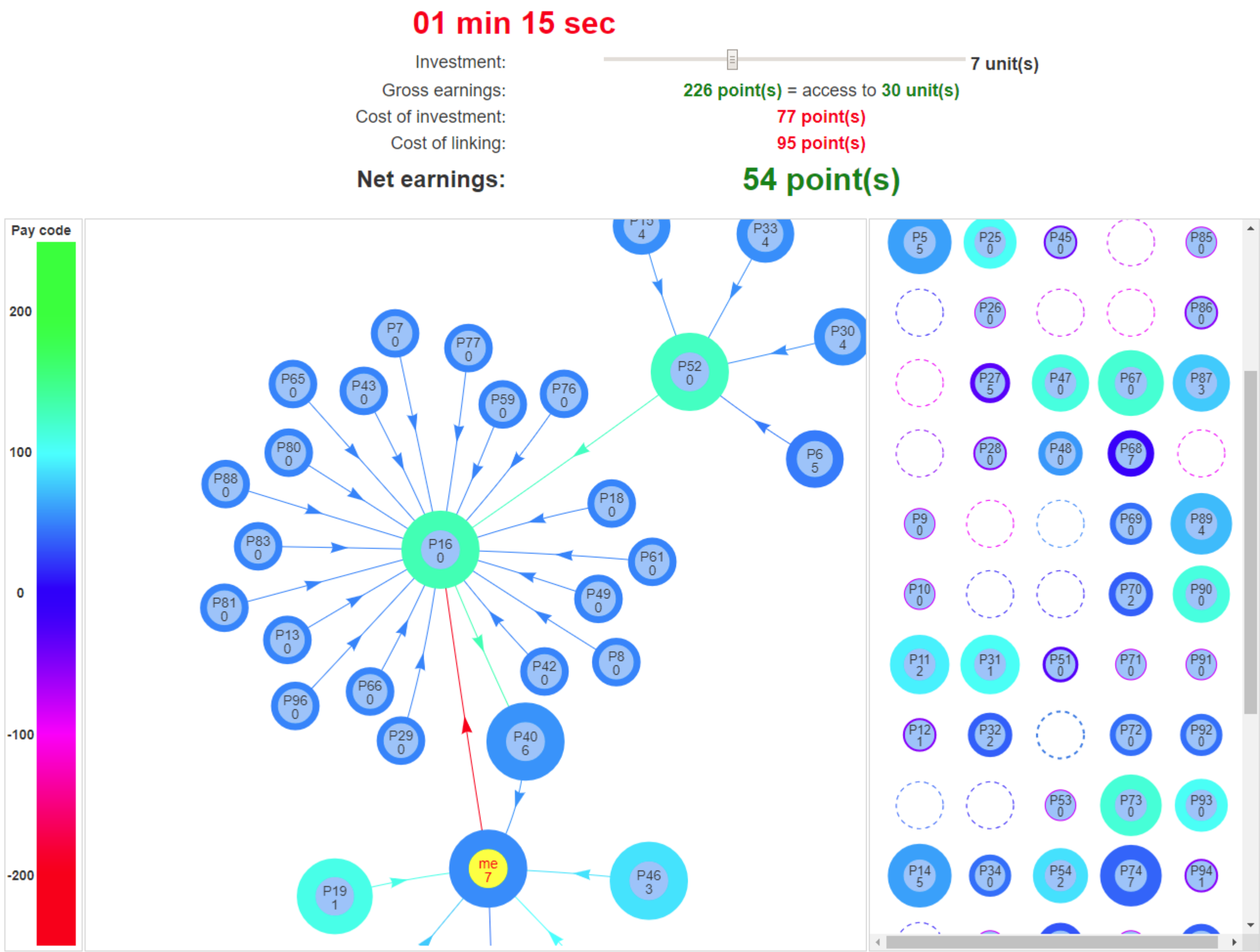


Figure 27: Example of decision screen: Connectors and Influencers experiment (PayInfo100 treatment)

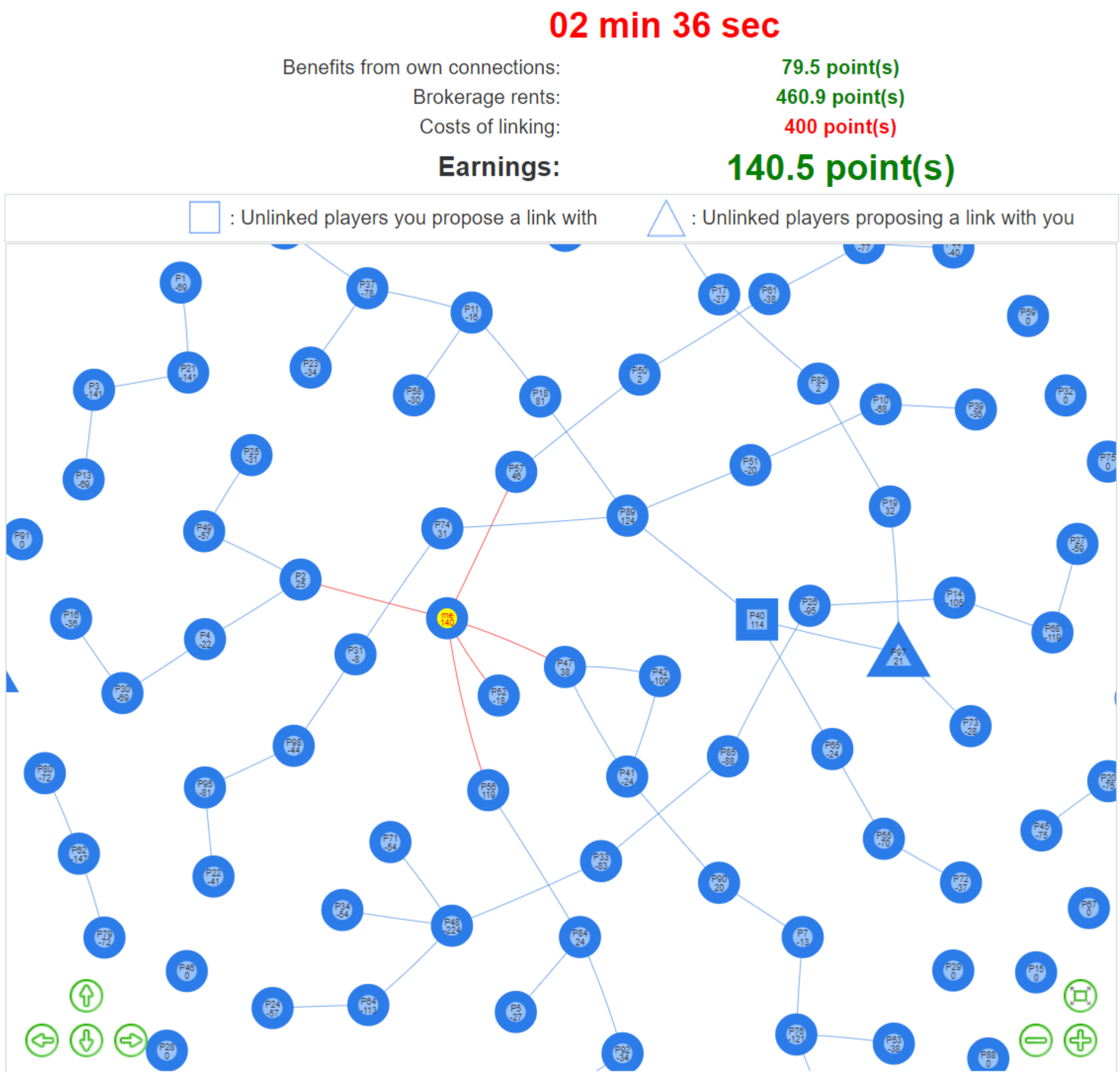


Figure 28: Example of decision screen: Brokerage and Market Power experiment (Criticality treatment, $N = 100$)

E Questionnaires

At the end of the experiment, subjects answered a set of surveys aiming at measuring various types of individual differences. More precisely, incentivized measures of comprehension in network game, social preferences, and risk preferences were used. Finally non incentivized personality measures were used before which subjects filled up a debriefing questionnaire that includes demographics information.

E.1 Comprehension check

In order to assess the subjects' comprehension of the network game played in each experiment, we provided a set of concrete questions, each of which with a unique correct answer. Each correct answer was rewarded with 0.1 euro for the subject.

E.1.1 Linking Game

The following first 2 questions were used across all treatments. Correct answers are "10 pts" to question 1, and "20 pts", "40 pts", or "80 pts" to question 2 depending on the treatment.

The third question depicted below was used in the treatment with $N = 50$ (the correct answer is "P1"). This question was adapted in all other treatments by matching the number of nodes to the group size in the experiment.

E.1.2 Connectors and Influencers

The following first 2 questions were used across all treatments (correct answers are "11 pts" to question 1, and "95 pts" to question 2).

The third question depicted below was used in the payoff information treatment with $N = 50$ (the correct answer is "P36"). This question was adapted in all other treatments by matching the number of nodes to the group size in the experiment, and by removing the colors in the baseline treatments.

The following questions 4 and 5 below were also used in the payoff information treatment with $N = 50$ (correct answers are "P1" for both questions 4 and 5). These questions were however adapted by again matching the number of nodes to the group size in the experiment. As before, these questions were also adapted to the baseline treatments by simply removing the colors to match the design of the actual game that subjects played.

Part 2/2: Questionnaire

Please answer the following questions. Any correct answer will earn you 0.1 euro.

Question 1: In the previous game, how many points did you receive from any of your neighbours?

- ○ 1 pts
- ○ 2 pts
- ○ 3 pts
- ○ 4 pts
- ○ 5 pts
- ○ 6 pts
- ○ 7 pts
- ○ 8 pts
- ○ 9 pts
- ○ 10 pts
- ○ more than 10 pts

Question 2: In the previous game, how many points did forming a link cost you?

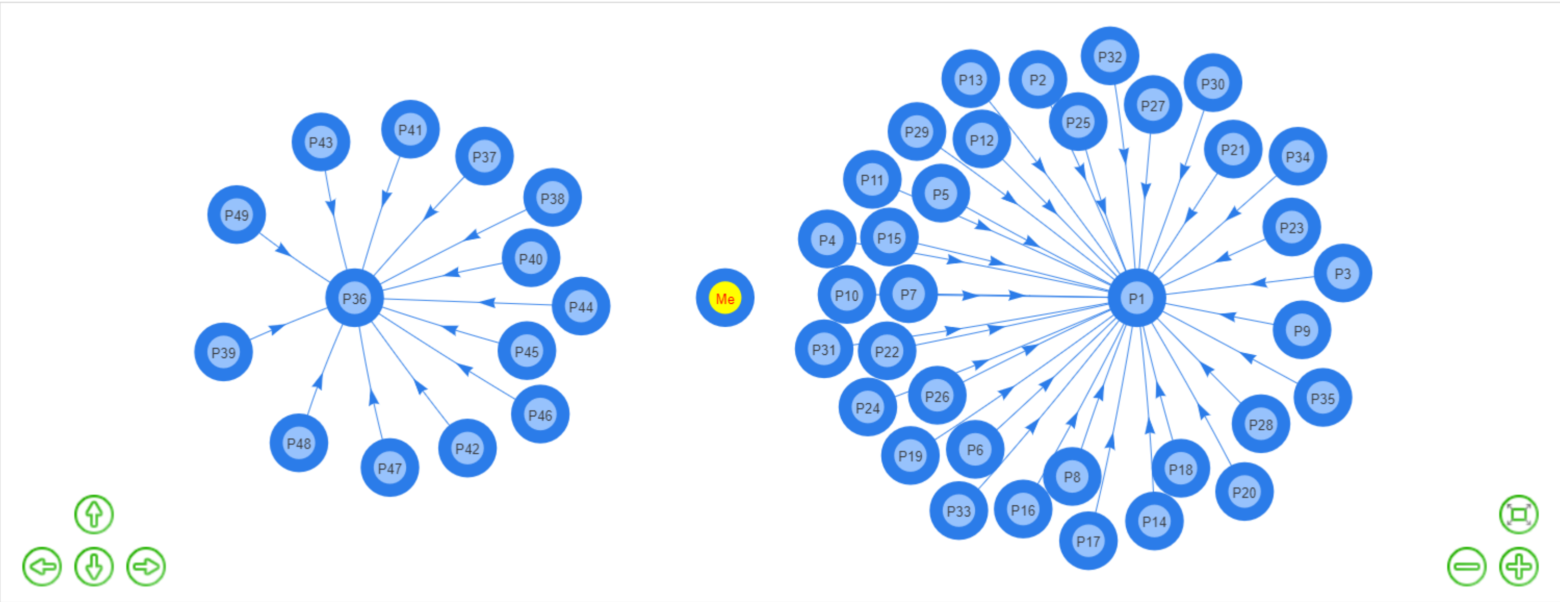
- ○ 0 pt
- ○ 20 pts
- ○ 40 pts
- ○ 60 pts
- ○ 80 pts
- ○ 100 pts
- ○ more than 100 pts

E.1.3 Brokerage and Market Power

All the same 6 questions were used across all treatments. The correct answers are "10 pts" to question 1 (in all treatments), "40 pts" or "more than 50 pts" to question 2 (depending on the treatments with group size $N > 10$), and "a randomly selected moment in the last 5 mins" to question 3. Note that in the case where $N = 10$, question 2 was adapted by substracting 2 pts from every options (the correct answer is then "8 pts").

The following questions 3, 4, and 5 relate to best response behavior in forming a link in some given network with $N = 50$. Correct answers are as follows: "P1" in question 4 (all treatments); "P1" (hub of the left hand side star) in question 5 (all treatments); "P18" (only node connecting the left and right component) for the Criticality treatment, and "P1" (center of wheel on left hand side) for the Betweenness treatment in question 6. Those questions were adapted to other treatments with different group sizes.

Question 3: In the hypothetical network below, please select one link (if any) that you think is most beneficial for you to form (remember that forming one link costs 100 points).
You may form at most one link by double clicking on the corresponding node. Click on Next to validate your answer.



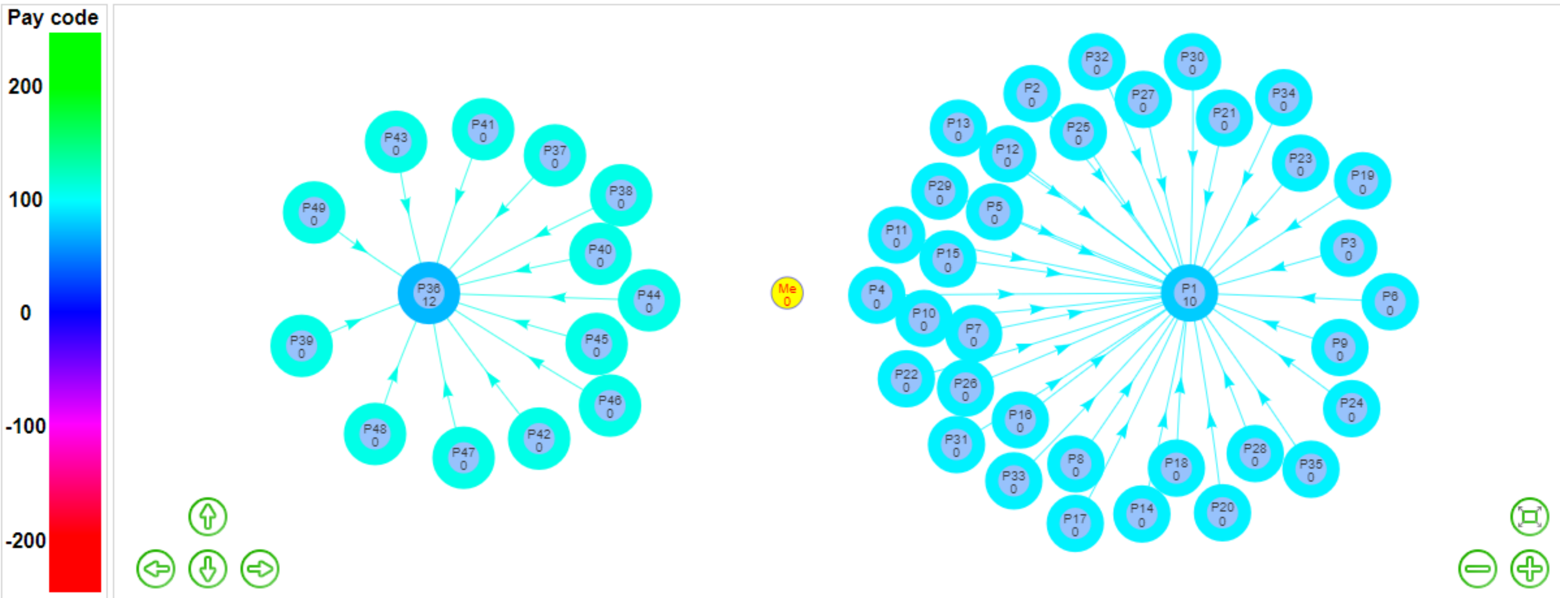
Question 1: In the previous game, how many points did investing one unit cost you?

- 1 pts
- 11 pts
- 21 pts
- 31 pts
- 41 pts
- 51 pts
- more than 51 pts

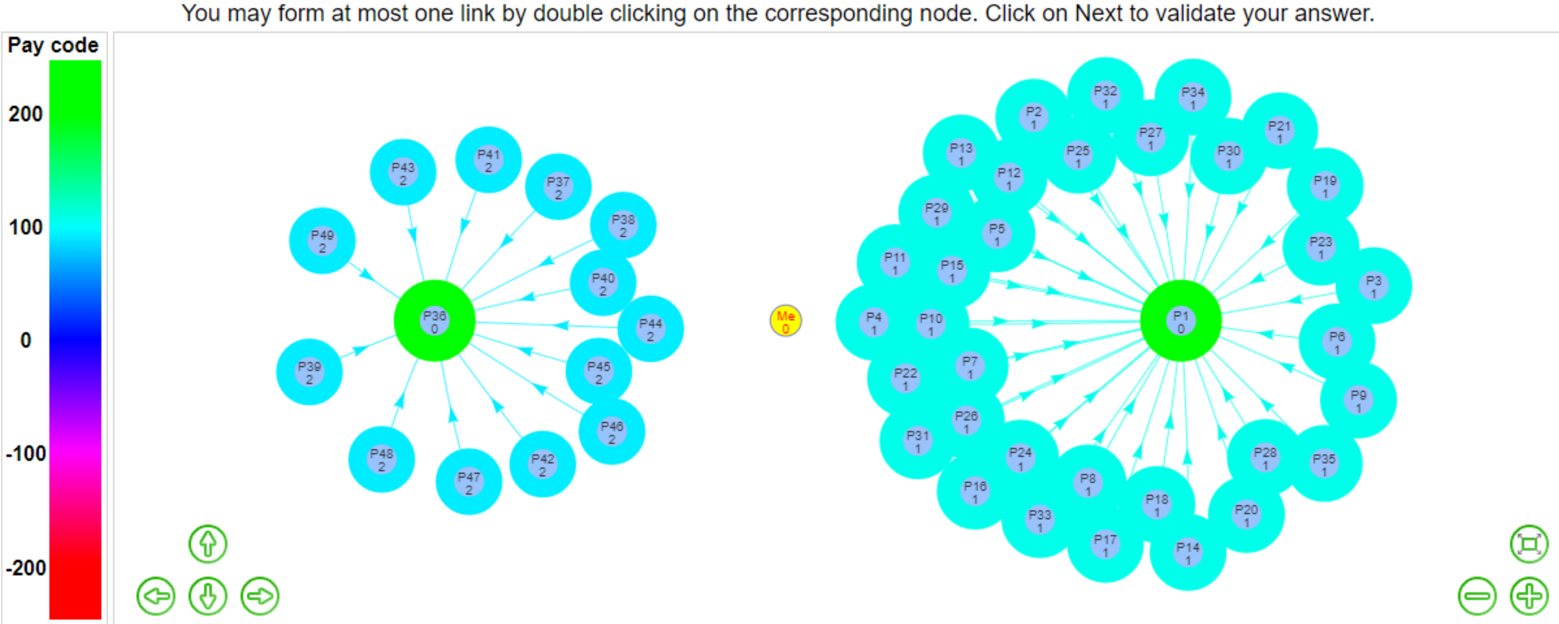
Question 2: In the previous game, how many points did forming a link cost you?

- 0 pt
- 25 pts
- 45 pts
- 65 pts
- 95 pts
- 115 pts
- more than 115 pts

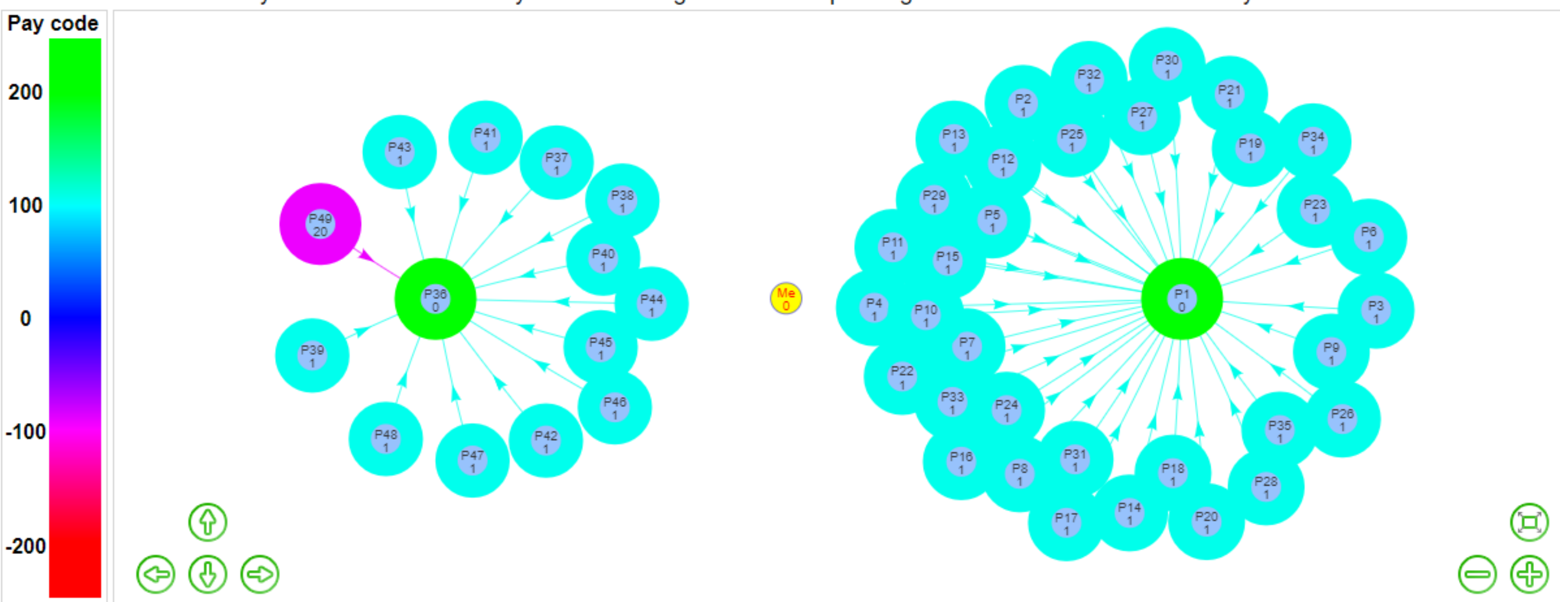
Question 3: In the hypothetical network below where you invest 0 unit, please select one link (if any) that you think is most beneficial for you to form (remember that forming one link costs 95 points).
You may form at most one link by double clicking on the corresponding node. Click on Next to validate your answer.



Question 4: In the hypothetical network below where you invest 0 unit, please select one link (if any) that you think is most beneficial for you to form (remember that forming one link costs 95 points).
You may form at most one link by double clicking on the corresponding node. Click on Next to validate your answer.



Question 5: In the hypothetical network below where you invest 0 unit, please select one link (if any) that you think is most beneficial for you to form (remember that forming one link costs 95 points).
You may form at most one link by double clicking on the corresponding node. Click on Next to validate your answer.



Part 2/2: Questionnaire

Please answer the following questions. Any correct answer will earn you 0.1 euro.

Question 1: In the previous game, how many points were created by a connected pair?

- 0 pts
- 2 pts
- 4 pts
- 6 pts
- 8 pts
- 10 pts
- more than 10 pts

Question 2: In the previous game, how many points did forming a link cost you?

- 0 pt
- 10 pts
- 20 pts
- 30 pts
- 40 pts
- 50 pts
- more than 50 pts

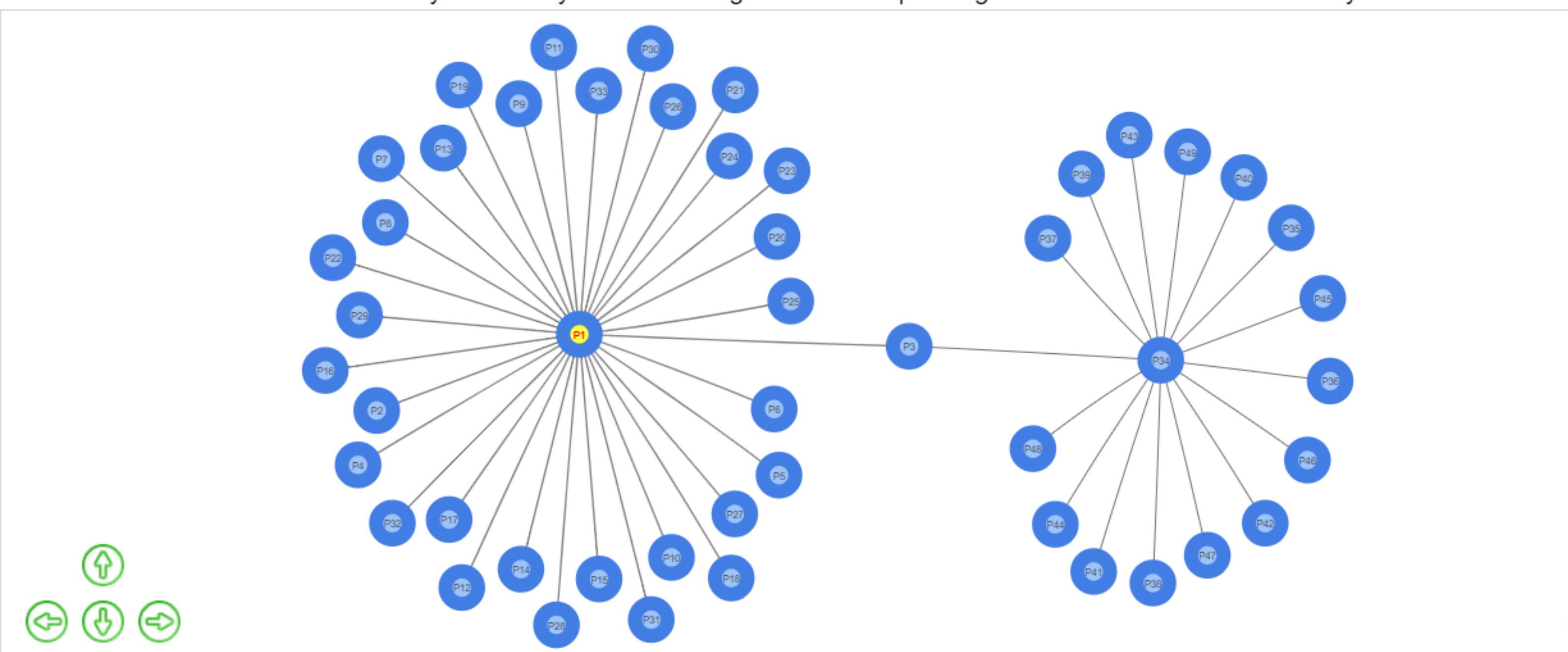
Question 3: In each of the previous games, your earnings were determined by considering everyone's choice at:

- A randomly selected moment in the 6 minutes
- A randomly selected moment in the last 5 minutes
- Every moment of the 6 minutes
- Every moment of the last 5 minutes
- The end of the 6 minutes
- The end of the 1st minute

Part 2/2: Questionnaire

Please answer the following questions. Any correct answer will earn you 0.1 euro.

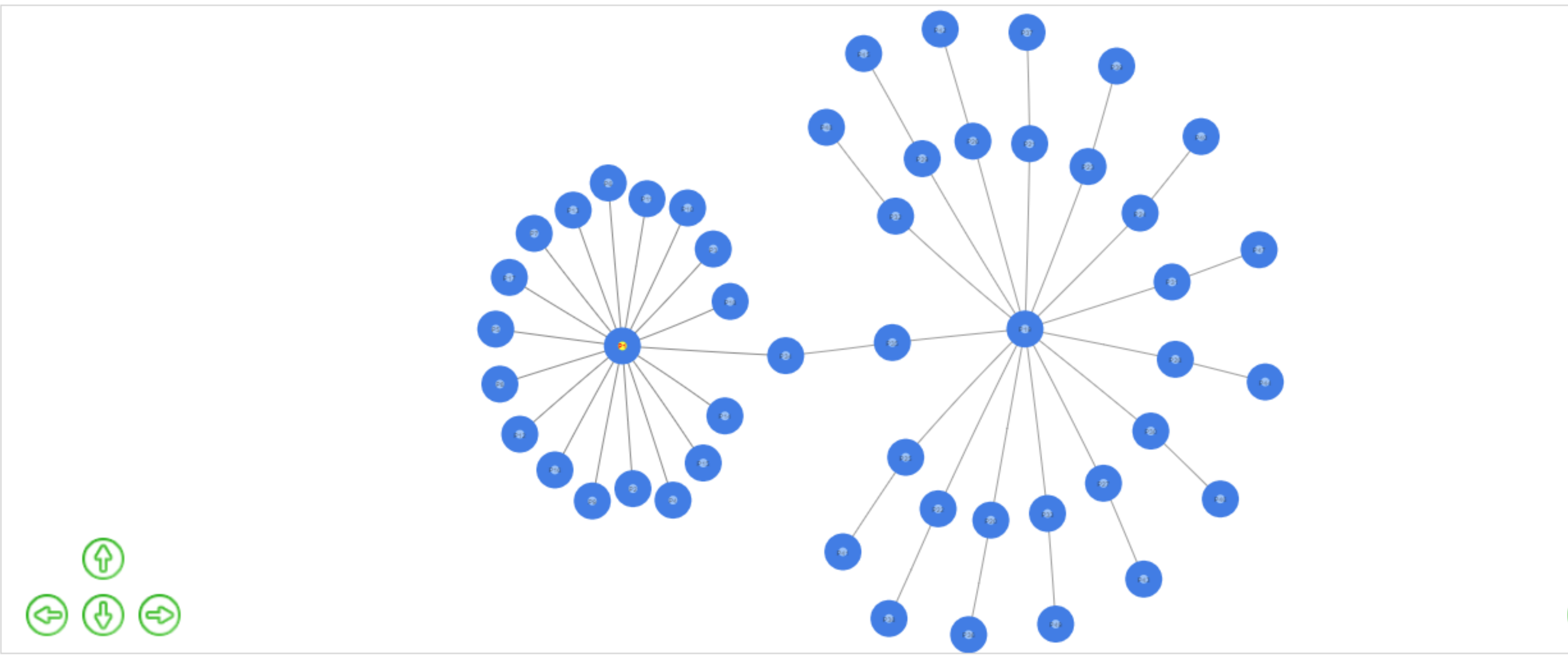
Question 4: In the hypothetical network below, assuming you are not linked with anyone, please select one node that you think is most ben you to link with.
You must form exactly one link by double clicking on the corresponding node. Click on Next to validate your answer.



Part 2/2: Questionnaire

Please answer the following questions. Any correct answer will earn you 0.1 euro.

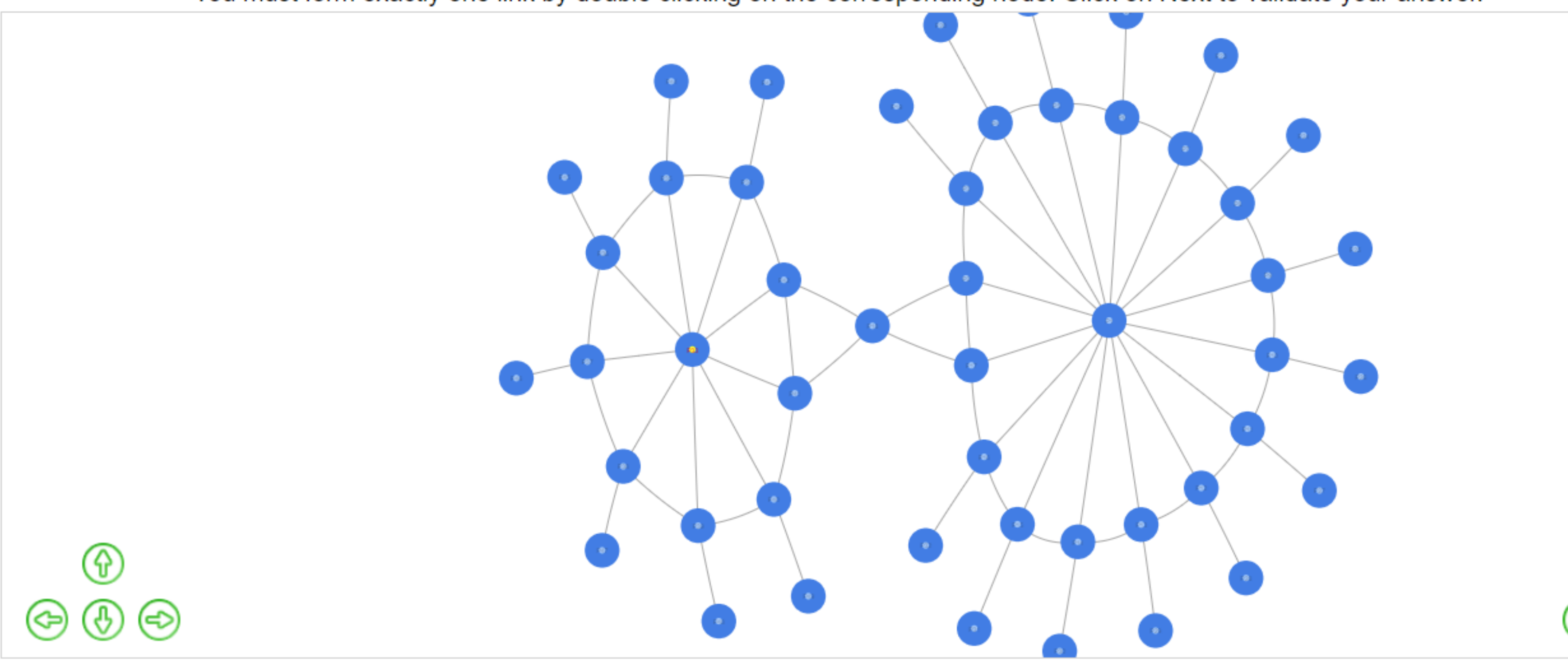
Question 5: In the hypothetical network below, assuming you are not linked with anyone, please select one node that you think is most ben
you to link with.
You must form exactly one link by double clicking on the corresponding node. Click on Next to validate your answer.



Part 2/2: Questionnaire

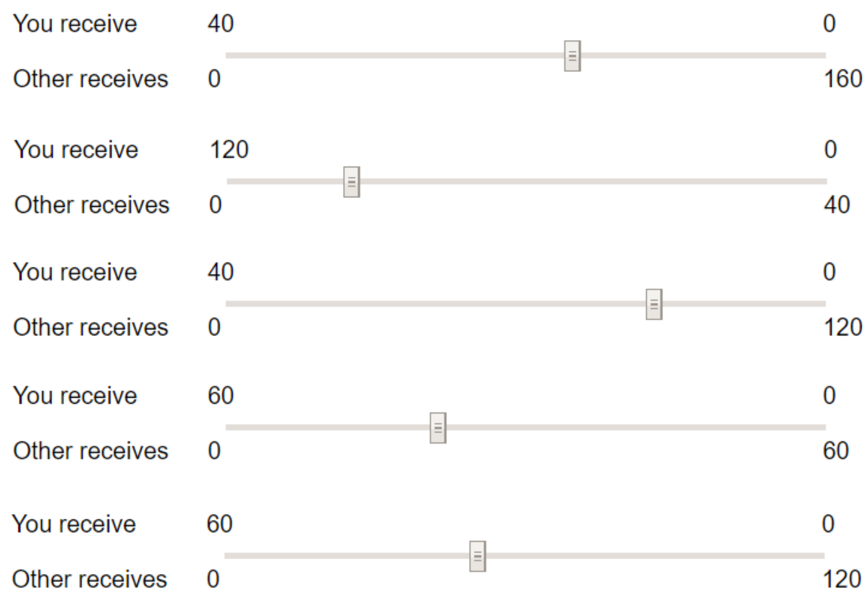
Please answer the following questions. Any correct answer will earn you 0.1 euro.

Question 6: In the hypothetical network below, assuming you are not linked with anyone, please select one node that you think is most bene you to link with.
You must form exactly one link by double clicking on the corresponding node. Click on Next to validate your answer.



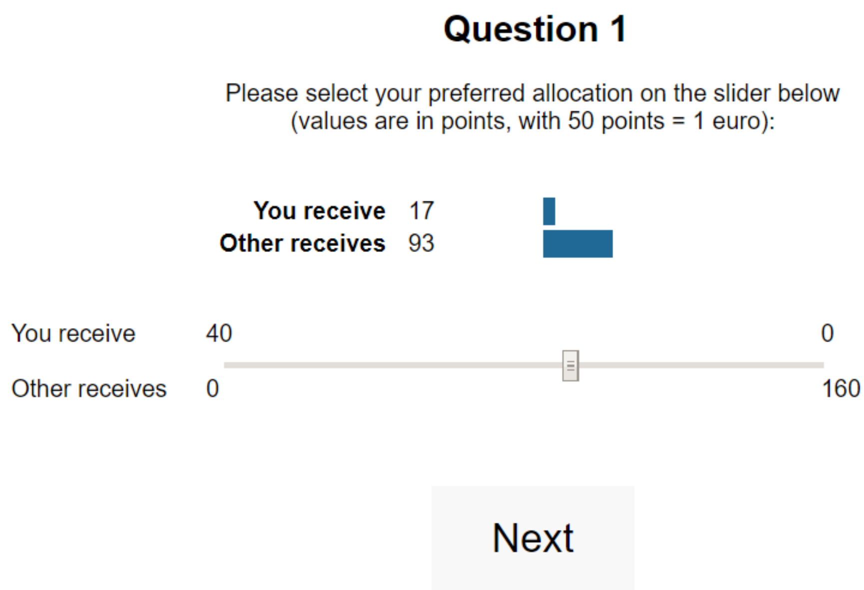
E.2 Social preferences

The social preferences measure was adapted from Andreoni and Miller [2002] and involved a series of five money allocation tasks between the decision maker and some anonymous external participants of another experiment at the LINEEX lab (corresponding payments were therefore made to these external passive participants). The five tasks used in our experiment were represented through sliders as shown in the following figure:



Note however that each question was presented in a different screen, and the order of presentation was randomized for every subject. Furthermore, 50 points were worth 1 euro both the subject, and the other anonymous external participant. Detailed instructions provided to the subjects, as well as a screenshot highlighting one of the above five questions are described below.

Instructions: *You are asked to answer a series of 5 questions, each of which consists of selecting an allocation of points that you most prefer between yourself and an anonymous randomly selected person who is participating to a different experiment in this lab. At the end of the study, we will randomly select your allocation for 1 of the 5 questions to determine the payments for both you and the other person in this part. Your decisions will remain unknown to the other persons you are matched with.*



E.3 Risk preferences

The risk preference measure was adapted from Holt and Laury [2002] and consisted of a series of five binary choices between lotteries, presented as in the figure below.

E.4 Personality test

Non incentivized measures were used through a simplified version of the Big Five personality inventory test adapted from Rammstedt and John [2007], as shown below.

You are now asked to make 5 independent choices between two lotteries. According to **Lottery A**, you can win 2.00€ with a certain probability **p**, and 1.60€ otherwise. According to **Lottery B**, you can instead win 3.85€ with the same probability **p**, and 0.10€ otherwise. For each of the following 5 choices, which only differ in the value of the probability **p**, please select the lottery that you prefer. At the end of the study, we will randomly select one of your 5 preferred lotteries to determine your payment in this question.

	Lottery A			**Lottery B**
Choice 1:	2.00€ with probability 20/100, 1.60€ with probability 80/100	○	○	3.85€ with probability 20/100, 0.10€ with probability 80/100
Choice 2:	2.00€ with probability 35/100, 1.60€ with probability 65/100	○	○	3.85€ with probability 35/100, 0.10€ with probability 65/100
Choice 3:	2.00€ with probability 50/100, 1.60€ with probability 50/100	○	○	3.85€ with probability 50/100, 0.10€ with probability 50/100
Choice 4:	2.00€ with probability 65/100, 1.60€ with probability 35/100	○	○	3.85€ with probability 65/100, 0.10€ with probability 35/100
Choice 5:	2.00€ with probability 80/100, 1.60€ with probability 20/100	○	○	3.85€ with probability 80/100, 0.10€ with probability 20/100

Next

How well do the following statements describe your personality?

I see myself as someone who...	Disagree strongly	Disagree a little	Neither agree nor disagree	Agree a little	Agree strongly
1. ... is reserved	○	○	○	○	○
2. ... is generally trusting	○	○	○	○	○
3. ... tends to be lazy	○	○	○	○	○
4. ... is relaxed, handles stress well	○	○	○	○	○
5. ... has few artistic interests	○	○	○	○	○
6. ... is outgoing, sociable	○	○	○	○	○
7. ... tends to find fault with others	○	○	○	○	○
8. ... does a thorough job	○	○	○	○	○
9. ... gets nervous easily	○	○	○	○	○
10. ... has an active imagination	○	○	○	○	○

Next

F Closeness centrality

In a given undirected network $\bar{g}$, the closeness centrality of a node i corresponds to $C_c(i, \bar{g}) = \frac{1}{n-1} \sum_{j \neq i} \frac{1}{d(i,j;\bar{g})}$. We then define the closeness centrality of the entire network $C_c(\bar{g})$ as follows:

$$C_c(\bar{g}) = \frac{\sum_{i=1}^{n} (C_c^{max}(\bar{g}) - C_c(i, \bar{g}))}{(n-2)(n-1)/(2n-2)} \tag{21}$$

Where $C_c^{max}(\bar{g}) = \max_i C_c(i, \bar{g})$. As a result, the closeness centrality of any network is bounded between 0 and 1. Note that this measure is maximized by the star network maximizes ($C_c(\bar{g}) = 1$) and minimized by the cycle, empty, and complete networks ($C_c(\bar{g}) = 0$).

G Regression Results

The regression analyses presented in the next sections rely on the aggregated data organized as follows. Given any round and group, subjects are classified in every second according to 3 types (according to their own indegree or number of received proposals): the most connected/popular individual, the 2nd most connected/popular individual, and the others. For each of those types, the data is then aggregated across the last 5 minutes of the round (e.g., mean outdegree, mean number of link proposals, mean effort, median payoff).

G.1 Linking Game experiment

	Mean outdegree		
	$N = 10$	$N = 50$	$N = 100$
Most connected	-0.173	4.598**	12.514**
	(0.143)	(0.800)	(3.453)
2nd most connected	-0.176***	2.373**	10.526***
	(0.028)	(0.669)	(0.924)
Mean of others	0.9348	1.3116	1.4820
Number of observations	60	60	60
R-squared	0.235	0.445	0.561

Notes: Robust standard errors, clustered by group, are reported in parenthesis. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively. All regressions include a constant, and dummies for groups and rounds.

Table 3: Regression analysis for the mean outdegree in the Linking Game experiment

	Median payoff		
	$N = 10$	$N = 50$	$N = 100$
Most connected	14.001***	-41.600	-1534.090**
	(4.479)	(78.058)	(673.610)
2nd most connected	4.699***	-48.320**	-2014.630***
	(0.714)	(23.790)	(462.112)
Mean of others	58.3514	270.1880	533.380
Number of observations	60	60	60
R-squared	0.2889	0.0476	0.2454

Notes: Robust standard errors, clustered by group, are reported in parenthesis. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively. All regressions include a constant, and dummies for groups and rounds.

Table 4: Regression analysis for the median payoff in the Linking Game experiment

G.2 Connectors and Influencers experiment

	Mean outdegree					
	Baseline			Payoff information		
	$N = 8$	$N = 50$	$N = 100$	$N = 8$	$N = 50$	$N = 100$
Most connected	-0.512**	0.164	1.417	-0.617***	0.426	5.448
	(0.152)	(0.457)	(1.674)	(0.030)	(0.848)	(5.832)
2nd most connected	-0.176**	-0.199	0.502	-0.184***	0.017	0.204
	(0.052)	(0.136)	(0.667)	(0.031)	(0.431)	(0.403)
Mean of others	0.7817	1.0425	1.1907	0.7280	1.0427	1.2462
Number of observations	60	60	45	60	60	45
R-squared	0.572	0.300	0.327	0.905	0.165	0.295

Notes: Robust standard errors, clustered by group, are reported in parenthesis. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively. All regressions include a constant, and dummies for groups and rounds.

Table 5: Regression analysis for the mean outdegree in the Connectors and Influencers experiment

	Mean effort					
	Baseline			Payoff information		
	$N = 8$	$N = 50$	$N = 100$	$N = 8$	$N = 50$	$N = 100$
Most connected	5.583***	13.046***	12.808***	5.350**	2.447	2.812
	(0.903)	(1.695)	(0.780)	(0.943)	(1.400)	(1.898)
2nd most connected	1.759***	11.483***	12.734***	2.469**	5.365**	2.154***
	(0.193)	(1.293)	(0.908)	(0.580)	(1.032)	(0.202)
Mean of others	2.7857	2.7602	3.4516	2.4311	2.2782	2.2477
Number of observations	60	60	45	60	60	45
R-squared	0.782	0.885	0.948	0.804	0.443	0.294

Notes: Robust standard errors, clustered by group, are reported in parenthesis. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively. All regressions include a constant, and dummies for groups and rounds.

Table 6: Regression analysis for the mean effort in the Connectors and Influencers experiment

	Median payoff					
	Baseline			Payoff information		
	$N = 8$	$N = 50$	$N = 100$	$N = 8$	$N = 50$	$N = 100$
Most connected	19.000***	-73.000***	10.750	25.750***	98.000***	114.167
	(2.422)	(6.339)	(24.164)	(7.629)	(17.009)	(287.472)
2nd most connected	17.000***	-81.500***	-87.000***	6.750*	-0.750	99.833***
	(3.013)	(8.989)	(6.780)	(3.940)	(18.832)	(22.450)
Mean of others	85.900	116.2875	146.1500	82.9875	98.1625	108.0833
Number of observations	60	60	45	60	60	45
R-squared	0.317	0.431	0.282	0.2149	0.1689	0.0868

Notes: Robust standard errors, clustered by group, are reported in parenthesis. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively. All regressions include a constant, and dummies for groups and rounds.

Table 7: Regression analysis for the median payoff in the Connectors and Influencers experiment

G.3 Brokerage and Market Power experiment

	Mean outdegree					
	Criticality			Betweenness		
	$N = 10$	$N = 50$	$N = 100$	$N = 10$	$N = 50$	$N = 100$
Most popular	0.302	4.949**	5.963	0.961	30.895***	58.925***
	(0.270)	(1.046)	(3.116)	(0.923)	(2.123)	(7.398)
2nd most popular	0.164	2.290	5.202***	-0.300	6.769*	28.389**
	(0.160)	(1.166)	(0.854)	(0.140)	(2.231)	(6.537)
Mean of others	3.9056	7.9036	9.3958	4.7242	10.4673	13.3495
Number of observations	60	60	60	60	60	60
R-squared	0.404	0.322	0.280	0.369	0.812	0.797

Notes: Robust standard errors, clustered by group, are reported in parenthesis. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively. All regressions include a constant, and dummies for groups and rounds.

Table 8: Regression analysis for the mean outdegree in the Brokerage and Market Power experiment

	Median payoff					
		Criticality			Betweenness	
	$N = 10$	$N = 50$	$N = 100$	$N = 10$	$N = 50$	$N = 100$
Most popular	-1.500***	-17.613**	-43.290	6.853*	1293.054***	2652.030***
	(0.542)	(8.266)	(49.025)	(3.858)	(197.464)	(250.957)
2nd most popular	-2.000***	-9.057	-29.590***	1.844***	20.100	53.030
	(0.529)	(7.649)	(10.043)	(0.565)	(19.252)	(96.143)
Mean of others	23.4708	132.0580	263.0945	18.0163	70.3097	105.3514
Number of observations	60	60	60	60	60	60
R-squared	0.2532	0.08	0.0874	0.1771	0.5752	0.598

Notes: Robust standard errors, clustered by group, are reported in parenthesis. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively. All regressions include a constant, and dummies for groups and rounds.

Table 9: Regression analysis for the median payoff in the Brokerage and Market Power experiment