UNIVERSITY OF CAMBRIDGE
Faculty of Economics

Institute for
New Economic Thinking

CAMBRIDGE WORKING PAPERS IN ECONOMICS
CAMBRIDGE-INET WORKING PAPERS

Connectors and Influencers

Syngjoo Choi
Seoul National University

Sanjeev Goyal
University of Cambridge

Frédéric Moisan
EM Lyon Business School

Abstract

We consider a setting in which individuals can purchase information at a cost and form costly links to access information purchased by others. The theory predicts that in every equilibrium of this game the network is a `star'. For small groups, there exists a unique purchase configuration – a pure influencer outcome, in which the hub node purchases information while all others free ride. For large groups, there exists, in addition, a pure connector outcome in which the hub purchases no information and the peripheral players purchase information. We test these predictions on a new experimental platform with asynchronous activity in continuous time.

We start with a baseline setting where subjects only see their own payoffs. We find that subjects create a star network. In small groups, the hub purchases equilibrium level information, but in large groups the hub purchases excessive information and as a result earns low payoffs. To study the reasons for this excessive investment we propose a treatment in which subjects see everyone's payoffs. We find that in small groups the pure influencer out-come obtains but that in large groups the pure-connector outcome now becomes common, suggesting that information and group size interact in powerful ways to shape networks and payoffs.

Reference Details

1935	Cambridge Working Papers in Economics
2019/09	Cambridge-INET Working Paper Series
Published	28 March 2019
Revised	17 April 2022
JEL Codes	C92, D83, D85, Z13
Websites	www.econ.cam.ac.uk/cwpe
	www.inet.econ.cam.ac.uk/working-papers

Connectors and Influencers

Syngjoo Choi[*] Sanjeev Goyal[†] Frédéric Moisan[‡]

April 17, 2022

We consider a setting in which individuals can purchase information at a cost and form costly links to access information purchased by others. The theory predicts that in every equilibrium of this game the network is a 'star'. For small groups, there exists a unique purchase configuration – a pure influencer outcome, in which the hub node purchases information while all others free ride. For large groups, there exists, in addition, a pure connector outcome in which the hub purchases no information and the peripheral players purchase information. We test these predictions on a new experimental platform with asynchronous activity in continuous time.

We start with a baseline setting where subjects only see their own payoffs. We find that subjects create a star network. In small groups, the hub purchases equilibrium level information, but in large groups the hub purchases excessive information and as a result earns low payoffs. To study the reasons for this excessive investment we propose a treatment in which subjects see everyone's payoffs. We find that in small groups the pure influencer outcome obtains but that in large groups the pure-connector outcome now becomes common, suggesting that information and group size interact in powerful ways to shape networks and payoffs.

JEL: C92, D83, D85, Z13.

[*]Seoul National University. Email: syngjooc@snu.ac.kr

[†]University of Cambridge. Email: sg472@cam.ac.uk

[‡]Emlyon Business School, GATE. Email: moisan@em-lyon.com

The paper was supported by the FET-Open Project IBSEN (contract no. 662725), the Keynes Fund and by Creative-Pioneering Researchers Program (Seoul National University), and was performed within the framework of the LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon. We thank participants at a number of seminars and conferences for very helpful comments.

1 Introduction

Large scale social networks are a defining feature of contemporary economy and society. Empirical research suggests that such networks exhibit a *law of the few*: the distribution of links is very unequal.[1] Given the social and economic implications of this inequality, it is important to understand the principles underlying the formation of these networks.

The economic approach to network formation takes the view that networks are created by individuals who compare the costs and benefits of linking. Beginning with the early work of Bala and Goyal (2000) and Jackson and Wolinsky (1996), this idea has been explored in a number of papers on network formation. A high level take away from this literature is that linking activity leads to unequal networks consistent with the 'law of the few'.[2] This result has been the subject of extended experimental investigation: the high level take away from this research is that subjects do not create networks that are in line with the theory (see e.g., Falk and Kosfeld (2012),Goeree, Riedl, and Ule (2009) and van Leeuwen, Offerman, and Schram (2019)). These experimental findings raise a question mark about the validity of an economic approach to understanding network formation.

A common feature of existing experiments is that the number of subjects is small (typically ranging between 4 and 8). Moreover, practically all the experiments require subjects to make simultaneous choices in discrete time. In a real world setting, groups are very large and individuals typically choose effort and linking at different points in time. The individual decision problem is complicated because the attractiveness of links depends on the efforts of individuals *and* also on the efforts by the neighbours of these individuals. As group size grows, these informational requirements become more demanding. So it is unclear if we can extend the findings from the small group experiments to more realistic settings. The work of Berninghaus, Ehrhart, and Ott (2006), Friedman and Oprea (2012) and Goyal, Rosenkranz, Weitzel, and Buskens (2017) suggests that continuous time experiments offer subjects more opportunities for choice and for learning and that they may offer better prospects for convergence to equilibrium than discrete time experiments. Our paper builds on this insight. We develop a new platform for network experiments in which individual choice is asynchronous and takes place in continuous time and also allows for

[1]See Barabási (2016) and Newman (2018).

[2]See e.g., Hojman and Szeidl (2008); Goyal (2022) and Bramoulle, Galeotti, and Rogers (2016) provide a survey of the large networks literature.

large groups of up to 100 subjects.[3]

This paper presents an experiment on a model taken from Galeotti and Goyal (2010) who study the following setting: Individuals personally acquire information and gather information through social contacts. Personal acquisition of information is costly and maintaining personal contacts also takes time and resources. Individuals therefore compare the relative costs of different sources of information. The theory predicts that for small groups there is a unique equilibrium with a pure influencer: the network is a star, the central node purchases information while all others connect to him and purchase no information. For large groups, there exist two equilibria: a pure influencer outcome and a pure connector outcome (in which the hub chooses zero information purchase while peripheral players purchase information).[4]

In the equilibria, any player can become the hub; in addition, in large groups, two very different purchase configurations can be sustained in equilibrium. Thus players face a challenging coordination problem and it is not clear if an equilibrium will be played and which of them will be selected. Efficiency and equity considerations are sometimes useful as guides to the behaviour of individuals and this may help in the selection of equilibrium.[5] In the model under consideration, the costs of effort and links are linear so the star network structure maximizes social welfare. Turning to equity, observe that the pure influencer equilibrium entails information purchase on the part of the hub and linking expenditures on the part of the spokes; this leads to a fairer distribution as compared to the pure connector outcome (in which the spokes do the information purchase as well as the linking and therefore earn much less than the hub). Correspondingly, the pure connector outcome offers large rents to the hub; and this creates an incentive for large information purchases in a bid to attract links. These considerations motivate a study of the effects of group size on behaviour: in line with existing literature, we study small groups of 4 and 8 (with a

[3]The platform is versatile and is being used to study questions relating to social learning in complex networks, strategic interaction on large networks, and the rise of dominant platforms. The concluding section below discusses this research.

[4]There is extensive empirical evidence for the existence of influencers and connectors; classical early studies of these phenomena include Katz and Lazarsfeld (2017). See Galeotti and Goyal (2010) and Goyal (2022) for an overview of the research in this field.
A number of paper have further theoretically explored the Galeotti and Goyal (2010) framework, see e.g., Baetz (2015), Perego and Yuksel (2016) and Herskovic and Ramos (2020). These models combine the two-way linking model of Bala and Goyal (2000) with the public goods model in networks model of Bramoulle and Kranton (2007).

[5]See for instance the seminal contributions of Charness and Rabin (2002), Bolton and Ockenfels (2000), and Fehr and Schmidt (1999).

unique pure influencer outcome) and large groups of 50 and 100 (with a pure influencer and pure connector equilibrium outcome).

We start with a baseline setting where subjects only see their own payoffs: in this case, for both small and large groups, subjects create star networks. In small groups, the hub purchases equilibrium level information, but in large groups the hub makes excessive information purchase and as a result earns low payoffs (Results 1 and 2). What are the reasons for this excessive investment by hubs in large groups? In a pure connector equilibrium the hub player earns large rents: so there is an incentive to make large information purchases in order to become a hub. It seems that subjects fail to anticipate that the benefits they can earn as a hub later do not compensate for the early costs of competing. This failure could be due to computational complexity: it is indeed very difficult to compute expected payoffs from being a hub in large groups as that depends on the linking and efforts of others. The only way for subjects to find out may be to actually reach that position, but by then it is too late to realize that the costs they incurred are not worth the benefits. This possibility leads us to a treatment in which subjects are shown the payoffs of everyone, including highly connected other individuals. We refer to this as the payoff information treatment.

Sharing everyone's payoff can help because a subject does not need to compute expected payoffs from being a hub, or wait until they reach that position to find out: they can simply observe how much others earn by reaching such a position. If subjects see that payoffs from being a hub are not that large they may compete less aggressively. This may alter their behaviour. This treatment can also test the non-material utility hypothesis because, should they care about status or efficiency or be altruistic, seeing others' payoffs should not have any effect on their behavior (if anything, it should reinforce it as they would see evidence that their extreme investments benefit everyone else in the group).

We find that in the payoff information treatment subjects create star networks in groups of all four sizes. This is in line with the theoretical prediction. In small groups, the hub continues to purchase information in line with equilibrium prediction. In large groups, however the hub makes small information purchases in the majority of the rounds. Indeed, in about 40% of the cases the hub purchases close to 0 information, giving rise to the pure connector equilibrium outcome! This is in sharp contrast to the pure influencer outcome observed in practically all the rounds of the baseline treatment.

In the final part of the paper, we study individual decision rules that account for the powerful effects of group size and information. We propose a parsimonious model that enables us to estimate both the likelihood of the occurrence of influencer and connector outcomes and the (contingent) decision rules on purchase activity. The decision rules incorporate the following ideas – that individuals compete to become a hub, that they choose a myopic best response, and that they imitate the activity level of the highest earning individual. Individuals place weight on these factors and choose their activity. We find that the influencer outcome is dominant in the baseline large group treatment, whereas the connector outcome is most likely in the payoff information large group treatment. In the influencer outcome state, individuals assign large weight to becoming the hub. In the connector outcome state, individuals place more weight on imitating the highest earning individual. The powerful interaction of group size and payoff information is thus explained by the differences in the occurrence of outcome states and the contingent decision rules.

Our paper contributes to the study of networks in economics. As we noted above, existing papers by Falk and Kosfeld (2012), Goeree, Riedl, and Ule (2009) and van Leeuwen, Offerman, and Schram (2019) find that experimental subjects reject the predictions of the model.[6] Our first contribution therefore is robust experimental evidence that subjects create hub-spoke networks, across group sizes and across payoff information treatments; this experiment offers support for an economic approach that explains networks in terms of the costs and benefits of links. Our second contribution is evidence on pure connector equilibrium, with large payoff inequality favoring the central hub node. To the best of our knowledge this is the first evidence of this interesting configuration in the literature. Our third contribution pertains to the role of group size and payoff information provision in shaping behaviour and payoff distributions.[7]

At a more general level, our paper contributes to the methodology of experiments by providing a new platform for conducting large scale experiments in continuous time (see e.g., Friedman and Oprea (2012) and Calford and Oprea (2017) and especially the early work of Berninghaus et al. (2006) on network formation). Existing studies are built on an

[6]We would also like to mention the experimental literature on games in networks (see e.g., Leider, Mobius, Rosenblat, and Do (2009), Charness, Corominas-Bosch, and Frechette (2007), 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 (see e.g., Riedl, Rohde, and Strobel (2016), Kearns, Judd, and Vorobeychik (2012)). Our experiment on the Galeotti and Goyal (2010) model supplements this latter strand of work. The novelty is that actions are asynchronous and in continuous time.

[7]In particular, our paper is closely related to the recent paper of van Leeuwen et al. (2019) who also test the model of Galeotti and Goyal (2010) model. There are a number of important differences in the specific models that are used and so the hypotheses tested are very different. Given these differences we discuss the relationship between the papers in section 5.1 below, after presenting our results.

experimental software, called ConG (Pettit, Friedman, Kephart, and Oprea (2014)) and have focused on small group interaction (see e.g., Friedman and Oprea (2012); Calford and Oprea (2017)). The novelty of our paper is that we develop an experimental software that is well suited for the study large group interaction. In order to overcome information overload of evolving networks and relax subjects' cognitive bounds in information processing, 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.

2 Theory

We present a model of linking and information purchases taken from Galeotti and Goyal (2010).

Let the set of players be denoted by $N = \{1, 2, \ldots, n\}$ with $n \geq 3$. Each player $i \in N$ simultaneously and independently chooses an effort level (that we will interpret as information purchase) $x_i \in \mathbf{R}$ 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 the set of linking strategies of player i be given by $G_i = \{0, 1\}^{n-1}$. 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. Observe that g is a directed graph; the closure of g 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 from efforts. Let $\eta_i(g) = |\{j \in N : g_{ij} = 1\}|$ be the number of links i has formed. 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. Define $N_i^l(\bar{g}) = \{j \in N : d(i, j; \bar{g}) = l\}$ to be set of players at distance l from i in $\bar{g}$.

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

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

where $c > 0$ 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 $a_l x_j$. 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 $f'(\hat{y}) = c$.

In what follows we would like to take into account direct and indirect flow of benefits and also decay in indirect benefits. A simple way to take these factors into account is to set $a_1 = 1$, $a_2 \in (0, 1)$, and $a_l = 0$, for all $l \geq 3$ in the payoffs equation (1). In this setting, building on arguments in Galeotti and Goyal (2010), we establish the following result for large costs of linking. The proof is provided in the online appendix.

Proposition 1. *Suppose payoffs are given by (1) that $a_1 = 1$, $a_2 \in (0, 1)$ and $a_l = 0$ for all $l \geq 3$. Then there exists a $\hat{k}$, such that for $k \in (\hat{k}, c\hat{y})$ the equilibrium network g^* is a periphery sponsored star and there exist two possible effort equilibrium configurations x^*:*

- *Pure influencer outcome (if $n \geq 3$): the hub invests $\hat{y}$ and everyone else invests 0.*
- *Pure connector outcome (if $n \geq 2 + \frac{k}{(c\hat{y}-k)a_2}$): the hub invests 0 and everyone else invests $\hat{y}/(1 + (n-2)a_2)$.*

We note that the assumptions $a_2 > 0$ and $n \geq 1 + \frac{k}{(c\hat{y}-k)a_2}$ are necessary for the pure connector equilibrium; if the condition on a_2 is not satisfied, then indirect access is of no value and following Galeotti and Goyal (2010) we can conclude that every (strict) equilibrium yields a pure influencer outcome. We also note that this assumption distinguishes our setting from the model considered by van Leeuwen, Offerman, and Schram (2019): they assume that $a_2 = 0$. So in their model, there only exists a pure influencer equilibrium. Assuming $a_2 > 0$ gives rise to a qualitatively different type of equilibrium that plays an important role in our experiment. However, this assumption is not sufficient to guarantee the existence of the pure connector equilibrium: if the group size n is too small, then periphery players making positive investments do not sufficiently benefit from each other, and therefore they can benefit from disconnecting the hub and making the optimal investment on their own.

Experimental parameters: We specify a functional form and parameters used in the experiment. The benefit function $f(.)$ is taken from Goyal, Rosenkranz, Weitzel, and

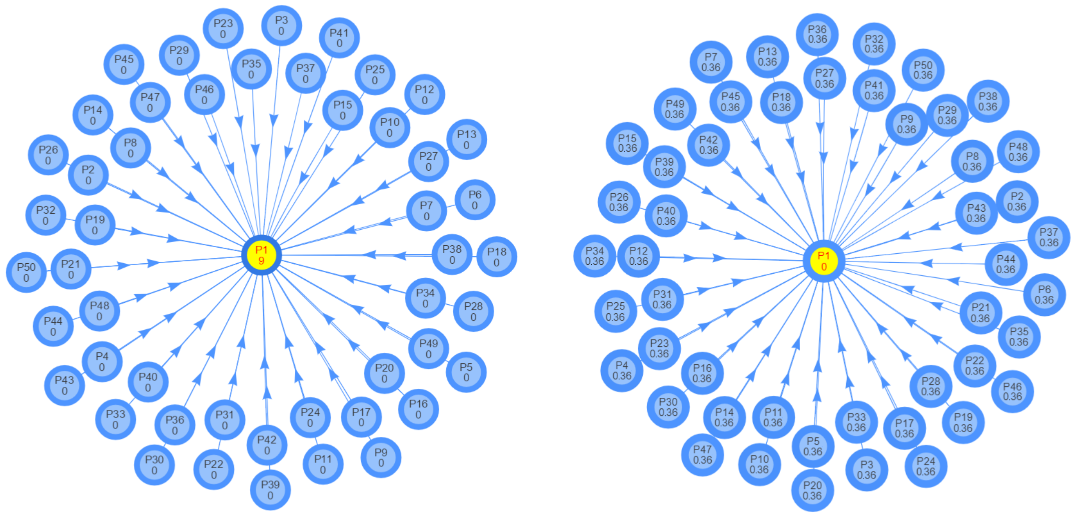


Figure 1: Pure influencer and pure connector equilibrium, $n = 50$ (individual effort levels are specified inside the nodes, below the node IDs)

Buskens (2017).

$$f(y) = \begin{cases} y(29-y) & \text{if } y \leq 14 \\ 196 + y & \text{else} \end{cases} \tag{2}$$

In our experiment, the cost of effort is $c = 11$, the cost of a link is $k = 95$ and the decay parameter $a_2 = 1/2$. Given these parameters, it can be checked that $\hat{y} = 9$ and the cutoff value of group size for the existence of the pure connector equilibrium is 49.5.

Proposition 1 tells us that in the small groups of 4 and 8 used in the experimental design a pure influencer is the unique equilibrium outcome; a single individual chooses 9, all other individuals choose 0 and form a link with this positive effort player. For large groups of 50 and 100 used in the experimental design, the network is a star with one hub, all the other players form a link with this hub, but two very different effort profiles can be sustained in equilibrium: the pure influencer outcome and the pure connector outcome. Figure 1 illustrates the pure influencer equilibrium and the pure connector equilibrium.

We note that in the experiment efforts take on integer values and that there is an upper bound on efforts, $\overline{x} = 20$; the set of efforts is given by $X = [0, 20]$. The predictions of Proposition 1 remain to hold as shown in Proposition 2 in the Appendix with the concept of ϵ-equilibrium.

For future reference, let us spell out the effort and payoffs in equilibrium under discrete values of effort. In the pure influencer equilibrium, the hub chooses effort 9 and the spokes choose 0 and form a link with the hub: the hub earns 81 and each of the spokes earn 85. In the pure connector ϵ-equilibrium, the hub chooses effort 0, eighteen or nineteen spokes choose 1 each, and the remaining spokes choose 0 each. In the presence of 18 active spokes, the hub earns 214, the active spokes earn 79.25, and the inactive spokes earn 85 (our interest is in the large payoff difference between spokes and the hub and these differences are of a similar order in the pure connector outcome with an inactive hub and 19 active spoke players).

2.1 Mapping theory on to the experimental design

The static model focuses on the trade-off between personal efforts and linking with others and reveals that individual incentives and strategic interaction lead to a network that has a very specific structure and that there are two effort configurations. To facilitate individual experimentation and learning, we consider a design in continuous time with asynchronous choice. However, this dynamic game opens the possibility of signalling, cheap talk, and reputation building, forces that go far beyond the original static game. The mapping from the static theory to the experiment is therefore not straightforward.

As we noted above in the introduction, this departure was partly motivated by considerations of learning opportunities. But it is perhaps worth making a higher level observation on the relation between theory and experimental design: our goal is to examine the economic implications of the trade-off between the costs of linking and the costs of personal investment as alternative routes to being well informed in a setting where information collected by different individuals are strategic substitutes. If the trade-offs identified in the theoretical model are central to the formation of unequal networks with great specialization then we believe that they should be reflected in choices made by subjects even if the experimental design departs on some dimensions from the static model. With this general observation in place, we now take up some more specific points.

First we note that we may consider the continuous time game as a sequence of simultaneous move games. In such an interpretation, it follows from standard arguments that a repetition of the static equilibrium constitutes an equilibrium of the game in the experiment. Second, we have a trial period of 60 seconds that has no direct pay-off relevance:

actions in this period may therefore be viewed as 'cheap talk'. This raises the question of whether cheap talk can select between different equilibria of the game. There is a large literature on this subject: a general message is that cheap talk is more likely to be effective in equilibrium selection if equilibria are Pareto ranked (see, e.g., Cooper et al. (1992), Farrell and Rabin (1996), Charness (2000)). In our setting, equilibria are not Pareto ranked, so, we believe that cheap talk is probably not an important factor in selecting equilibria in our analysis. The final remark is about the potential repeated game effects. From the work of Benoit and Krishna (1986), we know that repetition may be used to select among different stage game equilibrium and indeed even go beyond stage game equilibrium – to Pareto improving profiles of actions. This is certainly a possibility in our experimental design and indeed will be an important part of the experimental analysis.

With these observations in hand, we now state the hypotheses. Recall that the theory predicts a star network in both small groups and large groups. This suggests the following hypothesis with regard to network structure.

Hypothesis A *In both small and large groups the network is sparse, contains a highly connected hub and has small average distance.*

Next we state the hypothesis regarding information purchase. For small groups, the theory predicts the unique pure influencer outcome with the hub choosing 9 and every spoke choosing 0. On the other hand, in large groups there exists a pure influencer equilibrium and a pure connector equilibrium with the hub choosing 0 and many spokes choosing 0. Hence, the key potential effect of group size is therefore on the configuration of information purchase. This suggests the following hypothesis.

Hypothesis B *The hub in small groups chooses high effort and the hub in large groups chooses either high or low effort. Spokes in both small and large groups choose low effort.*

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 (for a proof of this property see Galeotti and Goyal (2010).) We have noted above in the introduction that the pure influencer is more equitable as compared to the pure connector equilibrium (and the computations in the previous section make this explicit). Following the experimental literature on the role of efficiency (Charness and Rabin (2002)) and inequality aversion

(Fehr and Schmidt (1999), Bolton and Ockenfels (2000)) this suggests that players would opt for the pure influencer outcome. On the other hand, the pure connector outcome offers the hub large rents and thereby creates an incentive for them to make large investments in a bid to become the hub. Although these considerations of efficiency and equity may suggest an equilibrium selection, it is ultimately an empirical question of whether to observe equilibrium outcomes and, if so, which outcome to observe in the experimental data.

Apart from group size that plays an important role in stating the hypotheses, as will be explained in the next section, we vary the visibility of others' payoffs in the experimental design. Motivated by the experimental literature of learning (see Camerer (2003) for a survey), this informational variation is hypothesized to play a further role in selecting which type of outcomes subjects coordinate on when both influencer and connector outcomes are predicted, that is, in large groups.

3 Experiment

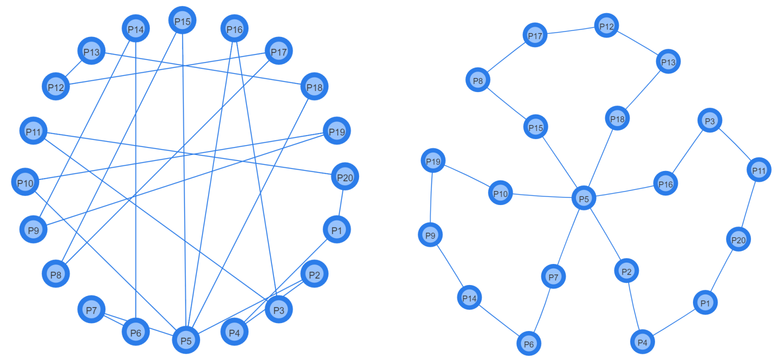
3.1 Challenges and methodology

As the complexity of subjects' decision making increases in scale, large-scale experiments on network formation pose several major challenges. This section discusses these challenges and points to ways in which our experimental software and design addresses each of them. Technical details are provided in the Appendix.

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. (2019)). 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. 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), and Jacomy et al. (2014)). Such algorithms are common, and have been previously used in Gallo and Yan (2015). The technical details of the specific algorithms are provided in the Appendix.

For example, consider a group of 20 people with fixed positions of nodes in a circle as depicted in Figure 2a; the exact network is barely perceptible by observing this figure (the same network visualized through a force-directed algorithm is shown in Figure 2b).

Visualization choices can potentially impact behaviour. In this paper we have opted



(a) Fixed visualization (b) Adaptive visualization

Figure 2: Examples of network visualization

to go with an approach that makes it easy for individuals to keep track of the evolving network. We are conscious that in actual practice, in the real world, individuals may not have access to such efficient visualization tools. The software we are developing allows for alternative visualization approaches and in future research we hope to examine the impact of visualization on behavior.

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. In view of the strategic complexity alluded to above, the issues of learning and behavioral convergence are particularly complicated. 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 near real time updating of all actions and linking by everyone.[8] At any instant of the game, every subject is free to asynchronously adjust their actions of efforts and linking.

Network information. In addition to the issue of network visualization, there is the issue of network information available to individual subjects. Our platform is flexible with respect to the level of information that is provided on the network. The platform allows us to show only local neighbourhoods to subjects and it allows us to go all the way to

[8] 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. Figure A7 in the Appendix illustrates the number of choices made by participants in our experiment.

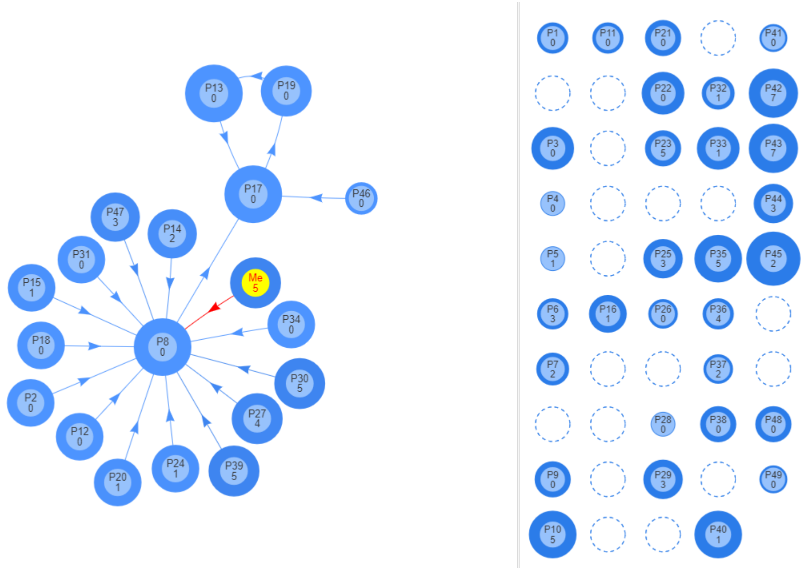


Figure 3: Network Information

showing subjects the complete network. In this paper, given a fixed network, for every subject, we can partition the subjects into two mutually exclusive subgroups: those who are located within (geodesic) distance 3 from the decision maker, and those who are located farther away. Figure A1 provides an illustration of network visualization and information in the experiment with 50 subjects. The left side of the screen shows the group of subjects within distance 3 of the decision maker identified through the yellow node as 'Me' (and all their links with other subjects within distance 3). The right side of the screen collects the subjects who lie at a distance greater than 3. Observe that in addition to local network information, subjects are informed about every subject's effort—presented as a number within the corresponding node along with that subject's ID. A node's access to others' efforts is reflected in the size of that node. In other words, a subject can, by comparing the number within a node to its size, make inferences on how much information/efforts it accesses from others.

3.2 Treatments and design details

We vary the group size $N \in \{4, 8, 50, 100\}$ and the visibility of others' payoff. Table 1 summarizes the 4×2 structure of our experiment.

		Group size			
		Small group		Large group	
		$N = 4$	$N = 8$	$N = 50$	$N = 100$
Others' payoff information	**NO**	**Baseline4**	**Baseline8**	**Baseline50**	**Baseline100**
	YES	**PayInfo4**	**PayInfo8**	**PayInfo50**	**PayInfo100**

Table 1: Experimental Treatments

At any instant in the 6 minutes game, a subject is free to asynchronously adjust their actions of efforts and linking. For the linking choice, the subject can form or 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 in more than 3 geodesic distance away), that subject along with his neighbors and neighbors' neighbors would be instantly 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.

During the experiment, each subject can also choose any level of effort by 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 net earnings (i.e., payoff $\Pi_i(x_g)$). Further information on the screen is provided in the Appendix.

3.3 Experimental procedures

The experiment was conducted at the Laboratory for Research in Experimental and Behavioral Economics (LINEEX) located in University of Valencia and at the Laboratory for Experimental Economics (LEE) that is located at the University Jaume I of Castellón. All the treatments except for $N = 100$ treatments were conducted at the LINEEX. The experimental sessions with $N = 100$ subjects were conducted through an internet connection between LINEEX and LEE (the number of subjects was then evenly distributed across the two locations). Subjects in the experiment were recruited from online recruitment systems of the two laboratories. A subject participated in only one of the experimental sessions.

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 a web browser (Google Chrome) on computer terminals and the experimental software was programmed using HTML, PHP, Javascript, and SQL. Sample instructions and interactive tutorials are available in the Appendix.

There were in total 18 sessions: 1 session of 16 subjects for each of the Baseline4 and PayInfo4 treatments, 1 session of 32 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. In each experimental session, subjects were (randomly) matched into a fixed group (if there were more than one group in a session) and interacted with the same subjects throughout the experiment. Therefore, there are 4 independent groups for each of the $N = 4$, $N = 8$, and $N = 50$ treatments and 3 independent groups for each of the $N = 100$ treatments. A total of 1096 subjects participated in the experiment.

The experiment consists of 6 rounds of the continuous-time game, each of which lasted for 6 minutes with the first minute as a trial period and the subsequent 5 minutes as the game with payment consequence. The first-minute trial period has two practical purposes: first, making subjects familiar with the play of the game and second, allowing them to build a network on their own, from an empty one to start with, for the game with real payment.[9] At the end of each round every subject was informed, using the same computer screen, of a moment randomly chosen for payment, detailed information on subjects' behavior at the chosen moment including a network structure and all subjects' efforts, and the resulting earning of the subject. While the membership of a group was fixed within a session, subjects' identification numbers were randomly reassigned at the beginning of every round in order to reduce potential reputation effects. The first round was a trial round with no payoff relevance and the subsequent 5 rounds were effective for subjects' earnings. In analyzing the data, we will focus on subjects' behavior and group outcomes from the last 5 rounds. At the beginning of the experiment, each subject was endowed with an initial

[9]The design with a non-paying trial period is used by leading papers in this literature such as e.g., Charness, Feri, Meléndez-Jiménez, and Sutter (2014). On the issue of generating a network, it is a common practice to assign an initial action profile in continuous-time game experiments as for example in Berninghaus et al. (2006) and Friedman and Oprea (2012).

balance of 500 points and added positive earnings to or subtracted negative earnings from that initial balance. Subjects' total earnings in the experiment amounted to 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).

At the end of the experiment, subjects took incentivized tasks to elicit social preferences and risk preferences. They are a modified version of the tasks proposed by 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 experimental game, and a debriefing questionnaire including demographic information. More details about these facts can be found in the Appendix.

4 Results: Baseline Treatments

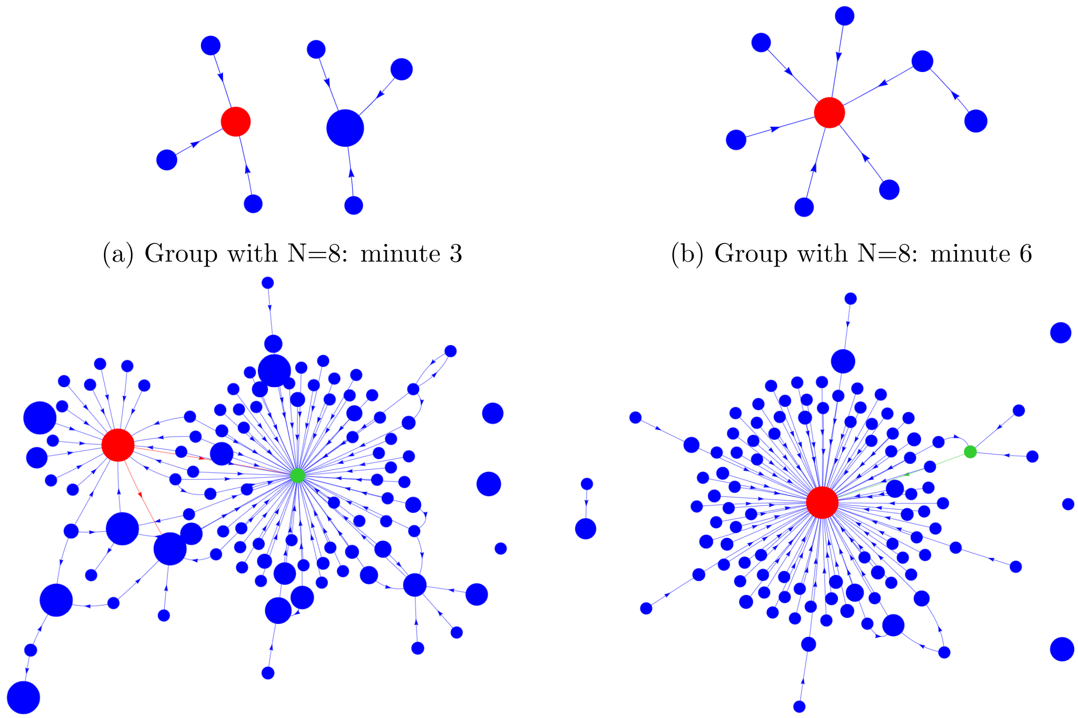
In all the data analyses that follow, the data used from every round of the game consists of 360 observations (snapshots of every subject's choices in the group) selected at regular time intervals of one second. Although some information about choice dynamics between two time intervals may be lost, we consider the possible impact of such a simplification as negligible to our analyses. Unless stated otherwise, the analyses are focused on data from the last 5 minutes of each round of the game.

We begin our study of experimental findings by presenting an overview of the dynamics of linking and efforts in small and large groups. Figure 4 presents snapshots taken from the experiment with $N = 8$ and $N = 100$ subjects (node size correlates with the individual effort level).[10]

In the small group ($N = 8$), subjects rarely make an effort in excess of 9, in line with the theoretical prediction. Over time, a single subject (red node in Figures 4a and 4b) emerges as the main hub with effort 9 while others make little investment (typically 0, as predicted by the theory).

By contrast, in the large group ($N = 100$), early in the game, a single subject (green node) emerges as the main hub with the maximum effort 20. There are other subjects who

[10]Full details of the 6 minute animations corresponding to Figures 4 and 8 are available at the following web site: `http://networks.econ.cam.ac.uk/net_formation/connectors_influencers.html`.



(a) Group with N=8: minute 3

(b) Group with N=8: minute 6

(c) Group with N=100: minute 3

(d) Group with N=100: minute 6

Figure 4: Snap shots from Baseline (N=8 and N=100)

make maximal effort (such as the red node). By minute 3, the green node player continues to be a hub but has substantially lowered her effort (see Figure 4c). Due to this shading of effort, she starts to lose some of her links to the red node player, who has kept her effort at maximum. The transition becomes clearer and at the end, the initial hub (the green node) has lost most of her links to the emerging hub (red node), as shown in Figure 4d. We now examine the data more systematically.

Let us first discuss the network structure. Figure 5 summarizes our findings about network structure across the different group sizes. We use average per capita degree as a measure of sparseness of a network[11], the maximum degree (normalized for group size) as a measure of being a hub, and average distance between two nodes as a measure of network

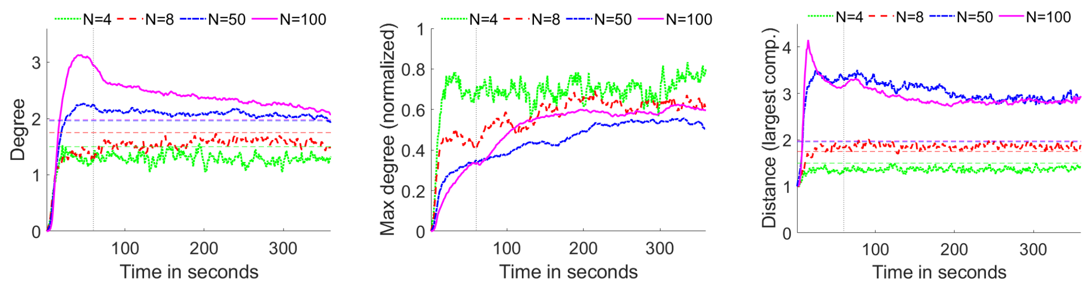
[11]Degree is defined as the number of links formed and received by an individual. This measure is justified here by the assumption in our model that information spillover is independent of the direction of links (see Section 2 for details). Note however that restricting this measure to incoming links only shows similar patterns.

closeness. In the left and right panel of Figures 5, the equilibrium benchmark is provided in dashed horizontal lines with the colour corresponding to group sizes. The equilibrium benchmark for normalized hub degree is equal to 1 for all group sizes and hence omitted.

First, we find that subjects create sparse graphs. Average degree is less than 2 in the small groups. In baseline50, the average degree is stable around 2. In Baseline100, average degree is falling over time to reach 2 at the end of the game. Recall that average degree is roughly equal to 2 in the star network (dotted horizontal lines in Figure 5a).

Second, the maximal (normalized) degree is high: it is between 0.6 and 0.8 in Baseline4, slightly above 0.6 in Baseline8, slightly below 0.6 in Baseline50 and between 0.50 and 0.60 in Baseline100.

Third, average distance is smaller than 2 in the small groups and it converges to 3 in the large groups.[12] Recall that the average distance (dotted horizontal lines in Figure 5c) in a star network would be close to 2. Therefore, we conclude that subjects create networks that are sparse, contain a hub, and have small average distance.



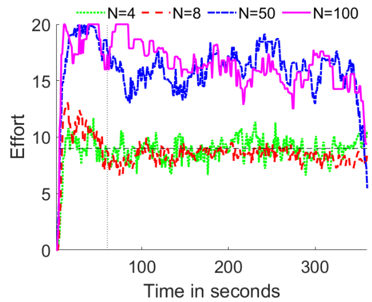
(a) Degree (b) Link inequality (c) Distance

Figure 5: Network Structure in the baseline treatments

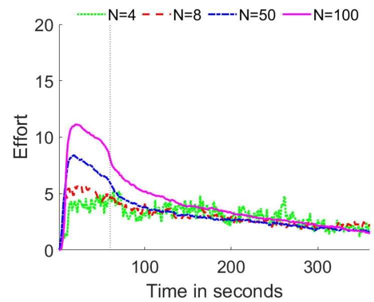
We summarize these findings about network structure.

Result 1 *In both small and large groups, subjects create a network that is sparse (average degree close to 2 in small groups and 3 in large groups), contains a hub (connected to 60%-70% of all nodes in small groups and 50%-60% of all nodes in large groups) and has small average distances (less than 2 in small groups and close to 3 in large groups).*

[12]Here we are considering the largest component, but the average size of the largest component is close to the group size in each treatment suggesting that the network is connected: average size of the largest component is 3.3 for Baseline4, 6.4 for Baseline8, 44.9 for Baseline50, and 94.8 for Baseline100.



(a) Most connected individual (hub)



(b) Others

Figure 6: Time series of efforts for the most connected subject and others in the baseline treatments

We conclude that the network structures in both small and large groups are in line with Hypothesis A.

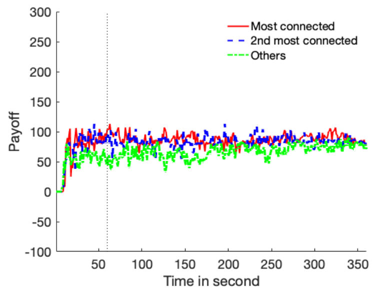
We next take up efforts. Figure 6 presents, for every second, the average efforts of the subject with the highest degree and the average efforts of all the other subjects.[13] In small groups the average effort by the highest degree subject is close to 9. There is a very sharp increase in effort by the most connected individual as we move from Baseline8 to Baseline50: it is 20 at some points in time and it remains above 15 for most of duration of the experiment. Similarly, in Baseline100 it remains above 13 for most of the period of the experiment. On the other hand, figure 6(b) shows that the average level of effort made by the other subjects is low and it steadily decreases over time.

We summarize these findings about the configuration of information purchase.

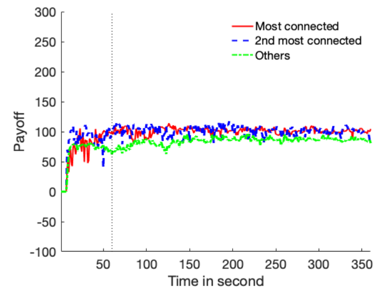
Result 2 *In small groups, the maximally connected subject chooses high effort (close to 9) and the spokes choose low effort (between 2 and 3). In large groups, the maximally connected subject chooses high effort (above 15 when N=50 and above 13 when N=100) and the spokes choose low effort (between 2 and 3).*

We conclude that in both small and large groups subjects' behavior is consistent with Hypothesis B: the hub chooses high efforts and the spokes choose low efforts. However, in

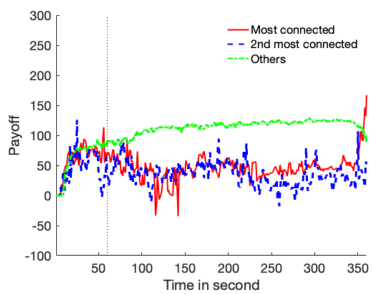
[13]Table A10 in the Appendix shows that the identity of the most connected individual differs across most rounds. In fact, we observe that 75% (67%) of subjects in a group of 100 (50) who become the most connected individual for at least 10 seconds in a round do so in only one of the five rounds. Overall, the fraction of subjects becoming the most connected individuals in any number of rounds is 7.6% (19.5%) in the group of 100 (50).



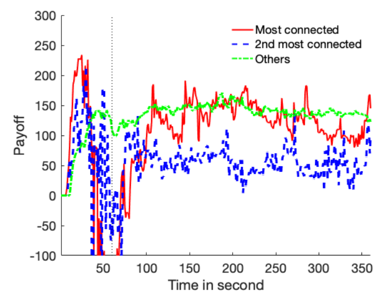
(a) Baseline4



(b) Baseline8



(c) Baseline50



(d) Baseline100

Figure 7: Time series of median payoffs for the three different types of subjects in the baseline treatments

large groups the hub chooses efforts far in excess of the equilibrium prediction. We examine the reasons for this sharp increase in efforts when we move from small to large groups next.

We start by asking if the excessive effort yields higher payoffs to the competing subjects. Recall that, in equilibrium, there is a negligible difference between the payoffs of the hub and spokes in the pure influencer outcomes but that the hub earns a much higher payoffs as compared to the spokes in the pure connector outcome. Figure 7 presents the dynamics of median payoffs obtained by three different types of subjects: the most connected and the second most connected individuals and the others. We see that the 2nd most connected subject in both Baseline50 and Baseline100 earns significantly lower payoffs than the 'other' subjects. The most connected subject in Baseline50 also earns lower payoffs than the 'other' subjects (except for the last 10 seconds). In Baseline100, they earn as much as the others for brief periods but the average payoffs are lower than others' payoffs. By contrast, in small groups, all three types of subjects earn similar payoffs.

We conduct a regression analysis of mean efforts made and a median regression of payoffs obtained by each type of subjects—most connected, 2nd most connected, and the others—on the dummy of large groups ($N = 50$ or 100). In this analysis, we define the types of subjects with the ranking of the fraction of time (across the five minutes) in which a subject is most connected.[14] The most connected individual is the subject who receives the most links for the largest fraction of time. The 2nd most connected individual is similarly defined. We refer to the rest of subjects as the 'others'.

Table 2 reports the regression results after controlling for round dummies, demographic information, comprehension test score, experimental measures of risk aversion and altruism, and personality. Robust standard errors (clustered by individual subject in the regression of efforts) are reported. Average efforts and median payoffs for each type of subjects in the small groups ($N = 4$ and 8) are also reported for comparison.

Table 2 says that there are significant group size effects on efforts and payoffs. The two most connected subjects make significantly higher efforts and earn substantially less in the large groups than in the small groups: 68% more efforts and 27% less payoff for the most connected subject, and 173% more efforts and 55% less payoff for the 2nd most connected subject.[15] Thanks to the intense competition of the two most connected subjects, the other subjects earned 44% more in the large groups than in the small groups.

To summarize: as the group size grows, individuals compete fiercely to become hubs. This leads them to invest very large amounts and, as a result, their earnings suffer. Indeed, in some cases the hubs actually make negative earnings. We observe that 25% (13%) of the most connected subject's sample in the Baseline50 (Baseline100) earn negative earnings (there is no incidence of negative earnings for the most connected subject in the small group treatments). We note that these negative earnings are neutralized by the initial endowment, which allows subjects to earn positive rewards at the end of the experiment despite having made losses in some round(s).

It is possible that in large groups, due to the complexity of the environment, some

[14] Figure A5 in the Appendix presents histograms showing the time fraction of different efforts over 5 minutes for the three different types of subjects across group sizes in the baseline treatment. The two most connected subjects in the large groups chose the maximum effort level, 20, for the majority of time, whereas in the small groups they chose significantly less with the mode of the most connected subject's effort being around the equilibrium effort level, 9

[15] Tables A1 and A2 in the Appendix report the replications of Table 2 by splitting the two large groups. The results remain similar with each of the large groups. In addition, we report the regression analysis of outdegree (the number of links) in Table A3 in the Appendix. We find that outdegree increases modestly for each type of subjects in the large groups: by about 1 for the most connected, by 0.8 for the 2nd most connected, and by 0.2 for the other subjects.

Table 2: Scale effects on effort and payoffs in the baseline treatments

	Mean effort			Median payoff		
	most connected	2nd most connected	others	most connected	2nd most connected	others
Large group	6.00***	9.04***	0.62*	-23.75*	-44.94**	37.12***
	(1.05)	(1.10)	(0.32)	(13.25)	(18.13)	(2.90)
Mean or median in small group	8.77	5.24	2.65	86.50	81.00	85.00
Number of observations	75	75	2590	75	75	2590
R-squared	0.59	0.64	0.04	0.19	0.23	0.08

Notes: Robust standard errors (clustered by individual subject in the regression analysis of efforts) are reported in parenthesis. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively. The dummy of large group assigns 1 if the group size is either 50 or 100, and 0 otherwise. All regressions include a constant, round dummies, age, female, education, comprehension test score, experimental measures of risk aversion and altruism, and Big 5 personality.

individuals who are keen to become hubs do not appreciate the payoff consequences of large efforts. If this is the case, then providing them information on everyone's payoffs would help them in choosing efforts more judiciously.

5 Payoff Information

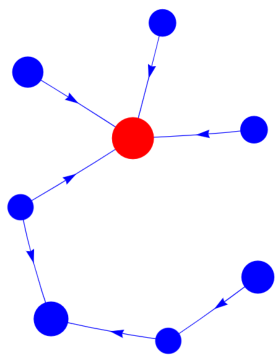
Clearly subjects need to be able to see their own payoffs in order to learn the profitability of different linking and effort combinations. In small groups, showing the payoffs of others may not be a first order issue, as subjects can compute these payoffs themselves in a fairly straightforward manner. However, in a dynamic continuous time game with fifty or 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. Therefore, in large groups the knowledge of others' payoffs may become a major factor. The first reason is learning dynamics: observing the others' payoffs, especially those who are hubs, could assist subjects in better appreciating the trade-offs associated with different courses of action. The second reason is fairness considerations: the two equilibria described in Proposition 1 exhibit very different level of payoff inequality across players. The pure-influencer equilibrium exhibits a minor payoff difference between the hub player and the spoke players, whereas the pure-connector equilibrium yields a much larger payoff difference between the hub player and the spokes players. These considerations motivate a treatment in which we provide information on others' payoff to everyone.

The literature of learning in games provides some guidance on the issue of information and learning, 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. In the recent body of network experiments (e.g., Goeree et al. (2009) and Falk and Kosfeld (2012)), researchers have tended not to show subjects the payoffs of others. However, when information on others' payoffs is available in particular in large groups where it is difficult to infer such information, subjects may follow a different behavioral rule. In fact, the experimental literature documents that human subjects may behave differently when information on the payoffs of other individuals is available (e.g., Huck et al. (1999)). With these observations in mind, let us describe how we proceed.

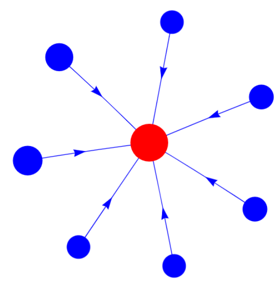
In the baseline treatments, subjects are shown their own payoffs but *not* others' payoffs. A subject is also shown the efforts and public good access for all other subjects, as shown in Figure A1 in the Appendix. In principle, therefore, a subject can infer the gross payoffs of any subject. But we believe that such inference would be challenging for subjects during a large scale continuous-time game, where the network and effort levels are evolving rapidly. To facilitate learning, we add information about every player's payoff through a set of colour codes as illustrated by Figure A2 in the Appendix. Specifically, the border of every node is coloured: the colour varies from green (high positive payoff) to red (high negative payoff). The scale of the colour code is presented at all times on the left hand side, as in Figure A2.

We start by presenting the dynamics of linking and efforts. Figure 8 presents snapshots taken in the payoff information treatment with $N = 8$ and $N = 100$ subjects. Observe that the properties of the network carry over as before. However, there is a change in the effort of the hub in the large group: the most connected individual (red node) starts at a high effort 14, but then gradually shades her efforts (reflected by her node size). By the end of the experiment, she is choosing an effort close to 0; as a result, the situation is close to a "pure connector outcome' (as in Figure 8d).

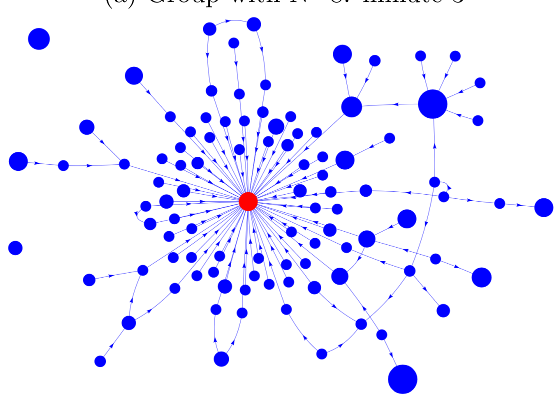
Figure 9 summarizes our findings about network structure in the payoff treatment. In the left and right panel of Figure 9, the equilibrium benchmark is provided in dashed horizontal lines with the colour corresponding to group sizes. The equilibrium benchmark



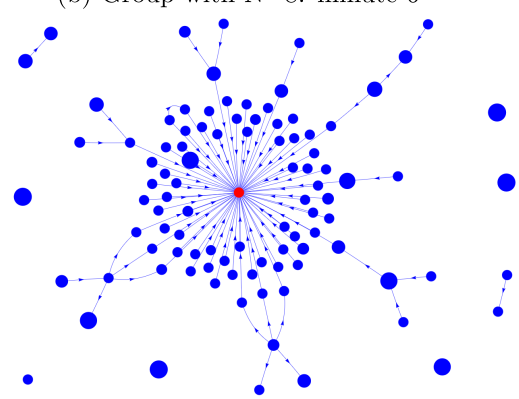
(a) Group with N=8: minute 3



(b) Group with N=8: minute 6



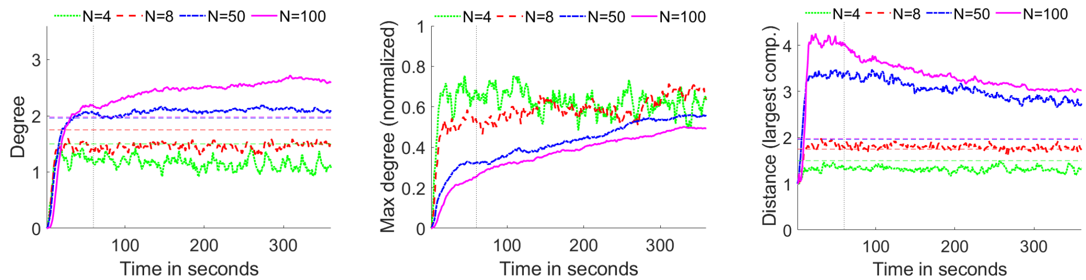
(c) Group with N=100: minute 3



(d) Group with N=100: minute 6

Figure 8: Snap shots with payoff information (N=8 and N=100)

for normalized hub degree is equal to 1 for all group sizes and hence omitted. We observe sparse networks across all group sizes (the average degree is less than 2 in small groups and it is less than 3 in large groups), the presence of a hub (the maximal (normalized) degree is 0.6 in small groups and 0.5 in large group), and small average distances (less than 2 in small groups and close to 3 in large groups).[16] These patterns are very similar to those observed in the Baseline treatments.



(a) Indegree (b) Link inequality (c) Distance

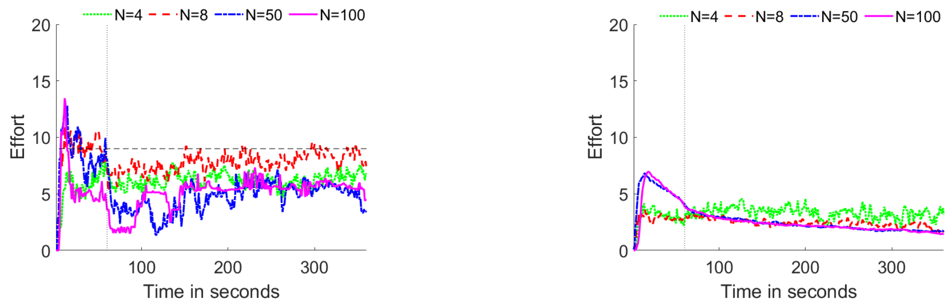
Figure 9: Network Structure in the payoff information treatments

Turning to efforts and their relation with degree, Figure 10 presents the dynamics of effort for the most connected individual and others over time. Compared to Figure 6 in the baseline treatments, we observe that behavior dynamics in the large group payoff information treatments are quite different: the efforts made by the most connected subjects are substantially lower. By contrast, in the small groups, the dynamics of efforts is similar across the payoff information treatment and the baseline. The behavior of 'other' subjects is similar across the two information treatments and across different group sizes.[17]

To better appreciate the effects of showing payoff information, it is helpful to separate the rounds into three types of outcomes with eyeballing the data: (1) a pure influencer outcome in which the hub chooses a large level of effort (in excess of 7) compared to efforts of others, (2) a pure connector outcome in which the hub chooses an effort less than 3, and (3) other outcomes in which the hub chooses effort between 3 and 7. In Section 6, we conduct a model-driven, statistical analysis of estimating the likelihood of influencer and connector outcomes with decision rules.

[16]The average size of the largest component is close to the group size in each treatment: 3.1 for PayInfo4, 6.1 for PayInfo8, 43.5 for PayInfo50, and 93.5 for PayInfo100.

[17]Table A4 in the Appendix presents the treatment effects of group size and information on efforts and payoffs, providing a statistical confirmation of the patterns arising from Figures 6 and 10.



(a) Most connected individual (hub)

(b) Others

Figure 10: Time series of efforts for the most connected subject and others in the payoff information treatments

Table 3 presents a classification of outcomes across treatments (together with the corresponding effort level by the hub and others averaged across the last 5 minutes).[18] Recall that there are 5 rounds per group, and 4 groups for sizes 4, 8 and 50 but 3 groups of size 100. So there are 20 rounds each for group sizes 4,8 and 50, and 15 rounds for group size 100 (for each of the Baseline and Payoff Information treatments). Table 3 tells us that in small groups there are practically no pure connector outcomes in the Baseline and the Payoff Information treatments. Matters are quite different when we turn to large groups. In the Baseline treatment, the pure connector outcome does not arise in a single round. However, in the Payoff Information treatment, the pure connector outcome arises in 9 rounds (out of 20) for group size 50 and in 6 rounds (out of 15) for group size 100. Thus in 15 out of 35 rounds the outcomes are consistent with the pure connector outcome where hub makes lower effort than others.[19]

Note that unreasonable choices made by a single individual can significantly contaminate the entire group's behavior. In our experiment, we occasionally observe that the most connected individual forms a disproportionate number of links (thereby generating large negative payoffs) for a significant period of time. We classify those individuals as outliers. In Table A9 from the Appendix, we therefore provide the same outcome classification that excludes rounds where the most connected individual earns less than -100 points during at least 60 seconds in the last 5 minutes. We observe that our results hold under this constraint.

[18]Time series of effort dynamics for the most connected individuals for each outcome types and each treatment are provided in Tables A5 and A6 in the Appendix.

[19]See Table A6 in the Appendix for an illustration of the corresponding time series analysis of efforts.

Treatment	N	Outcome Types Pure Influencer #Obs	 Hub	 Others	 Pure Connector #Obs	 Hub	 Others	 Other #Obs	 Hub	 Others
Baseline	4	14	10.2	3.6	0	.	.	6	5.8	3.3
	8	16	9.2	2.9	0	.	.	4	5.8	3.4
	50	19	16.3	3	0	.	.	1	5.8	3.2
	100	15	16.3	3.6	0	.	.	0	.	.
PayInfo	4	7	8.1	3.1	0	.	.	13	5.25	3.8
	8	16	8.4	2.6	0	.	.	4	5.3	3.5
	50	4	11.1	2.8	9	1.4	2.1	7	5.3	2.3
	100	4	11.9	2.1	6	0.6	2.4	5	4.9	2.2

Table 3: Mean effort in last 5 minutes across different types of Outcomes

Table 4 presents the regressions on treatment effects. In the small groups, there is little difference in each likelihood of infuencer and connector outcomes between the two information treatments. However, there is a marked difference between the two information treatments in large groups, regarding the occurrence of these outcomes. Thus, the key take away is the statistical significance of the interaction between payoff-information and group size. These observations are summarized as follows.

Result 3 *In small groups, subjects choose networks that are sparse, contain a hub and have small average distances. The hub chooses high effort and the spokes choose low effort.*

In large groups, subjects choose networks that are sparse, contain a hub and have small average distances. In some rounds, the hub chooses high effort while the spokes choose low effort, but in the majority of rounds the hub chooses low effort while the spokes choose higher effort.

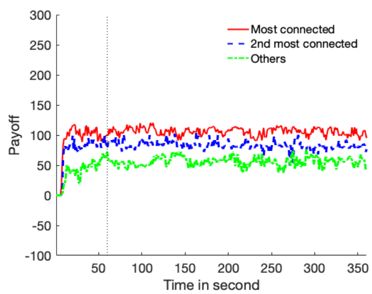
In small groups, subjects behave in line with Hypotheses A and B. In large groups, the networks created are in line with Hypothesis A. Regarding information purchase, in contrast to the baseline large groups, in the vast majority of the rounds in the payoff information large groups, the hub chooses low or very low efforts. This behavior has important payoff consequences for the most connected subjects who constantly earn more than others, in line with the pure connector prediction. This is clearly brought out in Figure 11.[20]

[20]Tables A7 and A8 in the Appendix presents time series of payoffs across different types of outcomes.

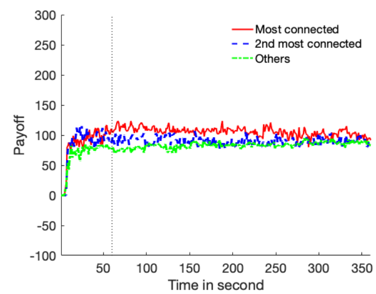
	Infuencer	Connector
Payoff info	-0.175	-0.000
	(0.142)	(0.018)
Large group	0.225**	0.004
	(0.094)	(0.020)
Payoff info × Large group	-0.568***	0.429***
	(0.176)	(0.141)
Frequency in small group baseline	0.750	0.000
Number of observations	150	150
R-squared	0.357	0.388

Table 4: Treatment effects on equilibrium selection

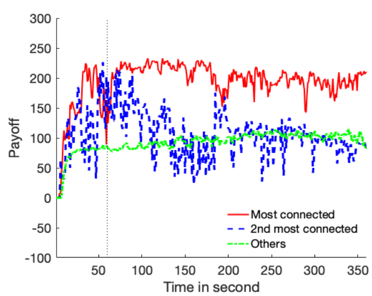
Notes: Robust standard errors clustered at the group level are reported in parenthesis. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively. The dummy of large group assigns 1 if the group size is either 50 or 100, and 0 otherwise. The dummy of Payoff info assigns 1 for the payoff information treatment. All regressions include a constant, round dummies, and group dummies.



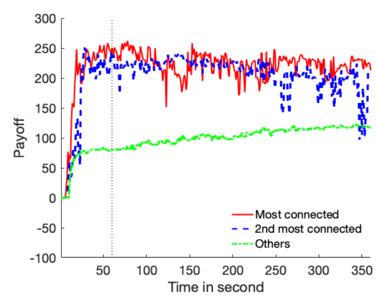
(a) PayInfo4



(b) PayInfo8



(c) PayInfo50



(d) PayInfo100

Figure 11: Time series of median payoffs for the three different types of subjects in the payoff information treatments

5.1 Relation with van Leeuwen, Offerman, and Schram (2019)

van Leeuwen et al. (2019) study the model of Galeotti and Goyal (2010) but set $\alpha_2 = 0$. With this restriction, the pure influencer is the unique equilibrium outcome. By contrast, we consider a variant of the model (that is perhaps slightly more natural) in which indirect benefits are positive, i.e., $\alpha_2 > 0$. As we showed in section 2, for small groups of 4 and 8, the pure influencer configuration constitutes the unique equilibrium profile in our setting too. On the other hand, with large groups of size 50 or 100, there exist two equilibrium profiles: a pure influencer and a pure connector outcome. This is the fundamental difference in the theoretical underpinnings of the two papers.

Turning to experimental findings, consider small groups of 4 and 8. Results 1 and 3 tell us that, both in the baseline as well as in the payoff information treatment, subjects abide by the pure influencer equilibrium. As there are no status-rents to be made in the pure influencer equilibrium, let us compare these findings with the no-status rents treatment in van Leeuwen, Offerman, and Schram (2019). They find that star networks and pure influencer outcome rarely arises (in less than 10% of the cases); a majority of the cases are instead characterized by mixing by players between low and high efforts. We believe that asynchronous activity in continuous time is probably the reason for the difference in findings.

Perhaps the biggest difference in the papers arises when we turn to large groups. They do not consider large groups, whereas most of our focus is in large groups. This is because in large groups, there exist two very different types of equilibria – the pure influencer and the pure connector. Our experiments with large groups present the first evidence for the rise of the pure connector outcome. Moreover, our analysis draws attention to the interaction between group size and provision of payoff information in shaping behaviour and creating large payoff inequalities.

As a final remark we note that the key finding of van Leeuwen, Offerman, and Schram (2019) is that subjects abide by a challenge-free equilibrium: this is a star network equilibrium of the repeated game in which the hub chooses high enough efforts to nullify any extra rents that she earns. Our evidence goes against this finding: there are very large payoff differences between the hub and the spoke in large groups (both in the baseline as well as in the payoff information treatment). In particular, in the baseline treatment, the hub makes excessive investments and earns much less than the spokes. On the other hand, in the payoff information treatment, the hub invests well below the spokes in 40% of the rounds and the outcome is close to a pure-connector outcome. The hub earns far more than the spokes – an outcome that is clearly not challenge-free. In our view, this brings out the powerful role of inertia: once a subject has emerged as a hub it is very difficult for a challenger to dislodge them without incurring large losses. Incumbency advantages appear to be much larger in the payoff information treatment as compared to the baseline treatment.

6 Estimating Treatment Effects on Decision Rules

We provide an estimable framework to gain a deeper understanding of the treatment effects on behavior and outcomes. As the differences in behavior across groups and information pertain mostly to the most connected individuals, we focus on them in this section. We examine the behavior of these two individuals in association with influencer and connector outcomes discussed in Tables 3 and 4.

Before introducing the framework, we first note that the behavior of most connected subjects in the large group baseline treatment is not justifiable from a dynamically rational point of view. A subject can guarantee herself an average payoff of 81 with an effort of 9 and zero links (regardless of what others do). Figure 7 shows that the two most connected subjects reach payoffs lower than this payoff in most cases. Therefore their behavior is not consistent with dynamic optimal choice.[21] This suggests that individuals in the baseline treatment who make large investments do not appreciate that their strategy is not as profitable as other strategies. As these excessive investments happen only in large groups, and it often disappears when payoff information for everyone is provided, we are led to the conclusion that increasing group size makes it difficult for highly connected subjects to keep track of the relation between actions and payoffs and this leads to over-investment. Therefore, we propose a single framework with selection of outcomes and boundedly rational decision rules that can account for the behavior of the two most connected subjects across the treatments.

Consider two most connected individuals, $i = 1, 2$, in a group. We assume, for the sake of parsimony, that they are either in the state of influencer outcome ($s = IO$) or in the state

[21]Figure A8 in the Appendix shows the continuation payoff that can be expected by one of the most connected individuals at any moment in time, by averaging the actual payoffs earned from that moment until the end of the game. In the large group baseline treatment, competing for the hub position is not profitable. On the other hand, under the payoff information treatment, it is more profitable in the large group as compared to the small group.

of connector outcome ($s = CO$). When they are in the state of influencer outcome, the two most connected individuals compete with the level of effort under influencer outcome, $\overline{x}$. When they are in the state of connector outcome, their decisions are based on the myopic best response effort. On top of this basic specification of decision rules in both outcomes, we allow the possibility that the individuals can follow the effort level chosen by a group member with the highest payoff in a previous time window. Allowing the possibility of imitating the higest earning individual in the decision rules enables us to examine the potential impacts of payoff information on the behavior of the two most connected individuals and group outcomes. The incorporation of 'imitate the best' to our specification of decision rules is partly motivated by the experimental literature of learning documenting that subjects adopt this imitation rule when they have necessary information to do so (see, e.g., Huck et al. (1999)). Because the computer screen was updated every 5 seconds or whenever the individual made a decision, we allow 3 seconds time lag in defining effort level in a previous time window.

Formally, we assume the following decision rule x_{it}^{d} with parameters, $(\overline{x}, \gamma, \lambda)$: for individual i in period t,

$$x_{it}^{d} = \begin{cases} \gamma \overline{x} + (1 - \gamma)\, x_{t-3}^{max} & \text{if} \quad s = IO \\ \lambda x_{it}^{mbr} + (1 - \lambda)\, x_{t-3}^{max} & \text{if} \quad s = CO \end{cases}$$

where x_{it}^{mbr} denotes the level of myopic best response effort for individual i at period t conditional on a game outcome at period $t-3$, and x_{t-3}^{max} denotes the level of effort chosen by an individual with the highest payoff in the group at period $t-3$. When the influencer outcome occurs, the decision rule places weight γ on $\overline{x}$ that captures the extent to which the two individuals compete for hub status. In the state of connector outcome, weight λ is placed on myopic best response effort.

Because the state of either influencer outcome or connector outcome is a latent variable, we estimate the probability of the game being in the state of influencer outcome, $Pr\{s = IO\}$. To use the maximum likelihood estimation method, we introduce an error term, ϵ_{it}, following the normal distribution with mean 0 and variance σ^2 independently and identically across i and t, in the decision rule: we treat an observed effort for individual i at period t, x_{it}, is a realization of a random variable, $h(x_{it}^{d} + \epsilon_{it})$, where $h\,(z)$ reports a value from the action set $\{0, 1, 2, ..., 20\}$ that is nearest to z. This stochastic formulation of decisions leads us to compute the likelihood of observing x_{it} conditional on being in the state of influencer outcome or connector outcome, $\psi\left(x_{it}|s=IO\right)$ and $\psi\left(x_{it}|s=CO\right)$.

Let $\left\{\{(x_{it}^{g},x_{it}^{mbr,g})_{i=1,2},x_{t}^{max,g}\}_{t=61}^{360}\right\}_{g}$ denote the samples of the two most connected individuals across group g over all the treatments. We estimate the parameters $(Pr\{s=IO\},\gamma,\lambda,\overline{x})$ for each treatment TR – the baseline small group treatment, the baseline large group treatment, the PayInfo small group treatment, and the PayInfo large group treatment – together with σ^2 common across all treatments. The log-likelihood function for the data is then given by

$$\mathcal{L}\left(\left\{\{(x_{it}^{g},x_{it}^{mbr,g})_{i=1,2},x_{t}^{max,g}\}_{t=61}^{360}\right\}_{g};\left\{Pr\{s=IO\}^{TR},\gamma^{TR},\lambda^{TR},\overline{x}^{TR}\right\}_{TR},\sigma^{2}\right)$$
$$=\sum_{TR}\sum_{i=1,2}\sum_{t=61}^{360}\ln\left[\Pr\{s=IO\}\,\psi\left(x_{it}|s=IO\right)+(1-\Pr\{s=IO\})\psi\left(x_{it}|s=CO\right)\right]$$

Table 5 reports the results that maximize the above-mentioned log-likelihood function. Standard errors, reported in parenthese, are computed with cluster bootstrapping at the level of game in every treatment with 500 replications. Several features of the results are noteworthy.

First of all, regarding the probability of being in the state of influencer or connector outcome, there is a marked difference between two information treatments in large groups. The estimated probability of being in the state of influencer outcome is 69% in the baseline large group treatment, whereas it is only 31% in the PayInfo large group treatment. In contrast, this probability is similar in the two information treatments in small groups; 55% in the baseline small group treatment and 60% in the PayInfo small group treatment. These findings about the likelihoods of influencer and connector outcomes corroborate those reported in Tables 3 and 4.

Second, we observe a significant effect of group size and payoff information on the decision rule. When they are in the state of influencer outcome, the two individuals in every treatment rarely pay attention to the effort level of a highest earning individual in the group and direct their full attention to seeking hub status. In addition, the estimated level of efforts under the state of influencer outcome show clear scale effects: they are much higher than the equilibrium level (around 18 in the baseline and 15 in the payoff information) in the large groups, whereas close to the equilibrium level (about 12 in the baseline and 8 in the payoff information) in the small groups.

When the two individuals are in the connector outcome state, we observe the powerful

		Influencer			Connector
Information	N	$Pr\{s = IO\}$	γ	$\overline{x}$	λ
Baseline	Small	0.55	0.88	11.66	0.49
		(0.031)	(0.003)	(0.234)	(0.062)
	Large	0.69	1.00	18.41	1.00
		(0.036)	(0.028)	(0.513)	(0.168)
PayInfo	Small	0.60	1.00	8.34	0.52
		(0.048)	(0.001)	(0.287)	(0.097)
	Large	0.31	1.00	14.64	0.00
		(0.040)	(0.100)	(1.906)	(0.135)
	σ^2		2.98		
			(0.125)		
	Log likelihood value		-265796.4		

Table 5: Maximum likelihood estimation

Notes: Standard errors, reported in parenthese, are computed by cluster bootstrapping at the game level with 500 replications.

interaction of group size and payoff information. In the baseline large group the two individuals place full weight on myopic best response effort, completely ignoring the effort level of a highest earning individual. In contrast, in the payoff information large group treatment, they imitate the effort level of a highest-payoff individual while completely ignoring their myopic best response effort. On the other hand, in the baseline small group and payoff information small group treatments, these two individuals adopt a similar decision rule that places equal weight on myopic best response effort and effort level of a highest-payoff individual.

These estimated decision rules deepen our understanding of the observed effort dynamics shown in Figures 6a and 10a. Excessive investment in the baseline large group treatment is driven by the high likelihood of influence outcome and competition for hub status through high efforts in that state. The low level of efforts in the payoff information large group treatment is explained by the high likelihood of connector outcome and the adoption of the imitate-the-best rule in that state.

7 Conclusion

There is a large body of research that describes the structure of large empirical networks. A recurring theme in this work is that networks exhibit great inequality in connections. The economic theory of network formation shows that the trade-off between the costs of linking and the benefits of direct and indirect links is resolved in strategic models in favor of unequal networks. However, experiments on these models show that subjects do not form such networks. This mismatch between the theory and the experimental evidence provides the motivation for our paper.

We develop a new platform for the study of network formation. The platform allows for continuous time linking and effort choice and it allows for large scale experiments. The paper presents an experiment on this platform with small groups of 4 and 8 and large groups of 50 and 100 subjects; we test the predictions on specialization on linking and efforts in the model of Galeotti and Goyal (2010). Our experiments provide strong support for the specialization prediction. Furthermore, our experiment offers first evidence for the emergence of a pure connector outcome. Finally, we show that group size and payoff information provision interact strongly to shape behaviour and payoffs.

The experiment reported in this paper is based on a network visualization technique of presenting subjects evolving networks in an efficient and visibly transparent manner. The reason of using this visualization is to facilitate subjects' learning about the evolution of networks during the games. Changing the visualization of networks may affect subjects' knowledge and perceptions about network changes and thus alter subjects' behavior. Exploring the effects of varying network visualization on game outcomes is an interesting avenue for future research.

The platform we have developed is versatile and can be used for the study of a wide range of questions. This is brought out in a series on companion papers. In Choi, Goyal, and Moisan (2020), we examine the formation of networks of intermediaries. The theory is permissive: equilibrium networks range all the way from hub-spoke networks with a dominant intermediary to the perfectly symmetric cycle network in which there are long chains of intermediation. In this experiment we study two-sided link formation (in contrast to the present paper where links are created unilaterally). In Choi, Goyal, Moisan, and To (2022) we study learning in large given canonical networks – Erdos-Renyi, Stochastic Block and Royal Family networks. The experimental design considers repeated discrete choice by subjects. Finally, in a new ongoing experiment, we study games on networks on large exogenous networks.

References

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. *Network science*. Cambridge university press, 2016.

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

J. Benoit and V. Krishna. Finitely repeated games. *Econometrica*, 53:905–922, 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. A theory of equity, reciprocity and competition. *America Economic Review*, 100:166–193, 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 goods in networks. *Journal of Economic Theory*, 135:478–494, 2007.

Y. Bramoulle, A. Galeotti, and B. Rogers. *Oxford Handbook on 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.

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.

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

G. Charness. Self-serving cheap talk: A test of aumann's conjecture. *Games and Economic Behavior*, 33(2):177–194, 2000.

G. Charness, M. Corominas-Bosch, and G. R. Frechette. Bargaining and network structure: An experiment. *Journal of Economic Theory*, 136:28–65, 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, S. Goyal, and F. Moisan. Brokerage. 2020.

S. Choi, S. Goyal, F. Moisan, and T. To. Learning in canonical networks. *Mimeo, University of Cambridge*, 2022.

R. Cooper, D. De Jong, R. Forsythe, and T. Ross. Forward induction in coordination games. *Economics Letters*, 40(2):167–172, 1992.

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

A. Falk and M. Kosfeld. Its all about connections: Evidence on network formation. *Review of Network Economics*, 11:Article 2, 2012.

J. Farrell and M. Rabin. Cheap talk. *Journal of Economic perspectives*, 10(3):103–118, 1996.

E. Fehr and K. M. Schmidt. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114:817–868, 1999.

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.

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

E. Gallo and C. Yan. The effects of reputational and social knowledge on cooperation. *Proceedings of the National Academy of Sciences*, 112(12):3647–3652, 2015.

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. *Networks: An Economics Approach.* MIT Press, 2022.

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.

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.

E. Katz and P. F. Lazarsfeld. *Personal influence: The part played by people in the flow of mass communications*. Routledge, 2017.

M. Kearns, S. Judd, and Y. Vorobeychik. Behavioral experiments on a network formation game. *ACM EC 2012*, 2012.

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. Newman. *Networks*. Oxford university press, 2018.

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.

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

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*, 2019.

Appendices: For Online Publication

A ONLINE: Appendix

A Theory

There are no general equilibrium characterization results available for this model. 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; this appears to be a natural case.

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 $k < c\hat{y}(n-2)a_2/(1+(n-2)a_2)$ or $n \geq 2 + \frac{k}{(c\hat{y}-k)a_2}$. 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 $a_1 = 1$ 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.

■

We next show that, under discrete values of personal effort, a sufficiently high cost of linking implies a pure influencer equilibrium (for any group size n) and a pure connector ϵ equilibrium (for a sufficiently large group size; e.g., with $\epsilon < 2$ for group size 50).

Proposition 2. *Suppose payoffs are given by (1), $a_1 = 1$, $a_2 \in (0, 1)$, $\hat{y} \in X = \{0, 1, 2, \ldots, \overline{x}\}$, and $c(\hat{y} + a_2 - 1) < k < c\hat{y}$. Every Nash equilibrium $s^* = (x^*, g^*)$ is such that g^* is a star.*

(a) For $n \geq 3$, there exists a pure influencer outcome: the hub chooses $\hat{y}$ and every spoke chooses 0.

(b) If $n \geq 2 + \frac{\hat{y}-1}{a_2}$, there also exists a pure connector ϵ-equilibrium, in which the hub chooses 0, m spokes invest 1 and the remaining other spokes invest 0.

Proof of Proposition 2: It follows from Proposition 1 that the pure influencer equilibrium always hold, regardless of n. Moreover, the pure connector equilibrium holds only if $n \geq 2 + \frac{k}{a_2(c\hat{y}-k)}$, in which case it requires every spoke to personally invest $\frac{\hat{y}}{1+(n-2)a_2}$. Since $c(\hat{y} + a_2 - 1) < k$ implies $\hat{y} < \frac{k}{c} + 1$, we have that $n \geq 2 + \frac{k}{a_2(c\hat{y}-k)} > 2 + \frac{\hat{y}-1}{a_2}$, and consequently $\frac{\hat{y}}{1+(n-2)a_2} < 1$ for any $n \geq 2 + \frac{\hat{y}-1}{a_2}$. However, since the lowest positive effort that can be made in the game is 1, there is a limited number of $m < n$ players who can benefit from making such minimum positive effort. In this case, each of those positive investors accesses $(m-1)a_2$ from forming a link and therefore earns $f(1+(m-1)a_2) - c - k$. They would earn $f(\hat{y}) - c\hat{y}$ should they form no link and invest $\hat{y}$. They would also earn $f((m-1)a_2) - k$, should they maintain their link and invest 0. The pure connector outcome is an ϵ-equilibrium whenever $\epsilon > f(\hat{y}) - f(1 + (m-1)a_2) + c(\hat{y} - 1) + k$ and

$\epsilon > f((m-1)a_2) - f(1+(m-1)a_2) + k$, where m is the number of investing spokes. Given the definition of f from (1), those conditions are satisfied for some values of $\epsilon \geq 0$ and $m \leq n$.

■

The combination of high linking costs and integer action levels leads to some complications with regard to the pure connector outcome that we now spell out. As noted above, a pure connector equilibrium does not exist for $n = 4$ or $n = 8$. In the treatments with 50 and 100 subjects there exists an ϵ-equilibrium in which 18 peripheral individuals choose 1 and the rest of the subjects choose 0. The equilibrium is 'approximate' because the periphery player who chooses effort 1 and forms a link with the hub earns 79.25 whereas she could earn 81 by deleting the link and instead choosing effort level 9. This asymmetry in behavior among the peripheral players makes the coordination problem more challenging than in the pure influencer outcome where it is solely about determining who is to be the hub.

B Experimental platform

A 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. (2019)). 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 subjects to learn their optimal choices, they must have a good idea of the evolving networks. An appropriate tool for visualizing networks 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), and Jacomy et al. (2014)).[22]

Clearly, different network environments will offer differ levels of transparency and information on network architecture. Our strategy is to start with a visualization approach

[22]Such algorithms are common, and have been previously used in Gallo and Yan (2015). The technical details of the specific algorithms are provided in the Appendix.

that is efficient and that allows us to systematically explore the effects of different variables – such as scale and variations in information on networks. Of course, the experimental platform is flexible enough to incorporate other ways of representing networks and can be used to explore the effects of network visualization itself on human behavior and network formation. Thus, the experiment reported in this paper could be interpreted as benchmark findings with efficient network visualization.

The network structure in Figure 2a from the Appendix can be represented in a transparent manner in Figure 2b with the network visualization tool we use. 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. Note that this feature does not aim at reproducing any cognitive representation of networks (i.e., how people mentally visualize networks), but instead attempts to facilitate access to information about the network structure, which people may use to make decisions. It is important to emphasize 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*.

The force-directed algorithms of the network visualization tool use attraction and repulsion forces 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 spatialisation 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):

$$F_x(o) = \frac{x_O - x_o}{dist(o,O)} F_g(o) + \sum_{o' \in N \setminus \{o\}} \frac{x_{o'} - x_o}{dist(o,o')} F_a(o,o') + \sum_{o'' \in N \setminus \{o\}} \frac{x_{o''} - x_o}{dist(o,o'')} F_r(o,o'') \quad (3)$$

$$F_y(o) = \frac{y_O - y_o}{dist(o,O)} F_g(o) + \sum_{o' \in N \setminus \{o\}} \frac{y_{o'} - y_o}{dist(o,o')} F_a(o,o') + \sum_{o'' \in N \setminus \{o\}} \frac{y_{o''} - y_o}{dist(o,o'')} F_r(o,o'') \quad (4)$$

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 at the following website: `http://networks.econ.cam.ac.uk/net_formation/connectors_influencers.html`.

We turn back to Figure 2 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 2b 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 (3) and (4) 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 2 and show how the algorithm makes the transition from the fixed visualization in Figure 2a to the adaptive visualization in Figure 2b. Six (slow-motion) snap shots of the transition are presented at the following website: `http://networks.econ.cam.ac.uk/net_formation/connectors_influencers.html`. 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 2b.

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*.

Model parameter setting used in the experiment:

- $K_g = -2000$
- $K_s = 0.04$
- $K_{cg} = 0.3$
- $L = 95$
- $D = 0.09$
- $T = 0.5$
- $V_{min} = 0.3$
- $V_{max} = 10$

Continuous time with asynchronous choices. We elaborate on some of the technical aspects of the platform in relation to these points. 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 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*. 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. **Information on Payoffs.** We turn finally to information on

payoffs: clearly subjects need to be able to see their own payoffs in order to learn the profitability of different linking and effort combinations. In the baseline treatments, subjects are shown their own payoffs but *not* others' payoffs. A subject is also shown the efforts and public good access for all other subjects, as shown in Figure A1. To facilitate learning, we add information about every player's payoff through a set of color codes as illustrated by Figure A2 in the Appendix. Specifically, the border of every node is coloured: the colour varies from green (high positive payoff) to red (high negative payoff). The scale of the colour code is presented at all times on the left hand side, as in Figure A2.

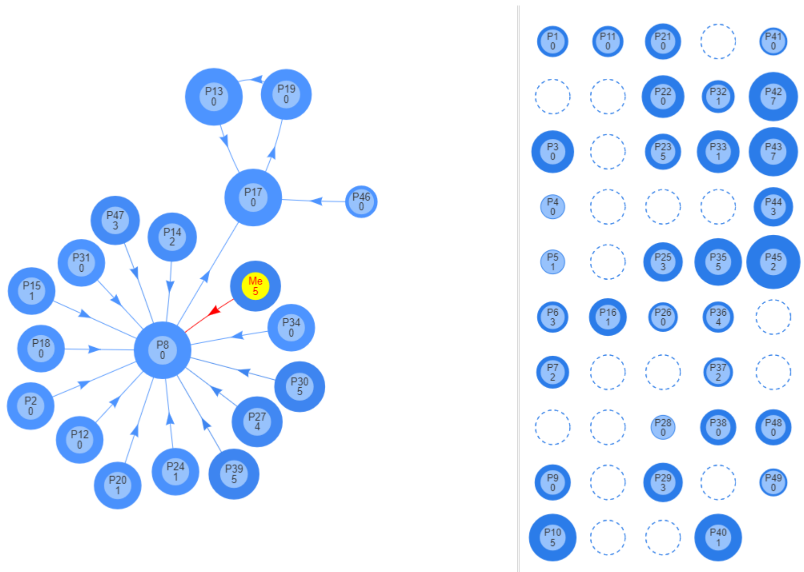


Figure A1: Network Information

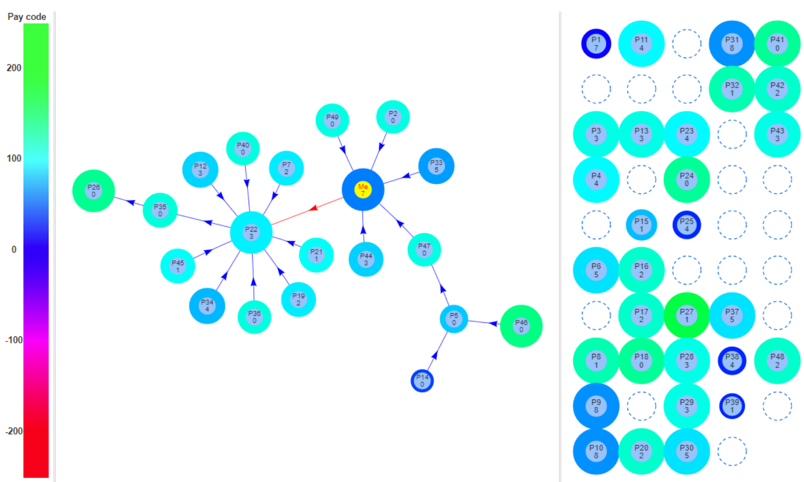


Figure A2: Payoff Information

C Experimental instructions

[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**.

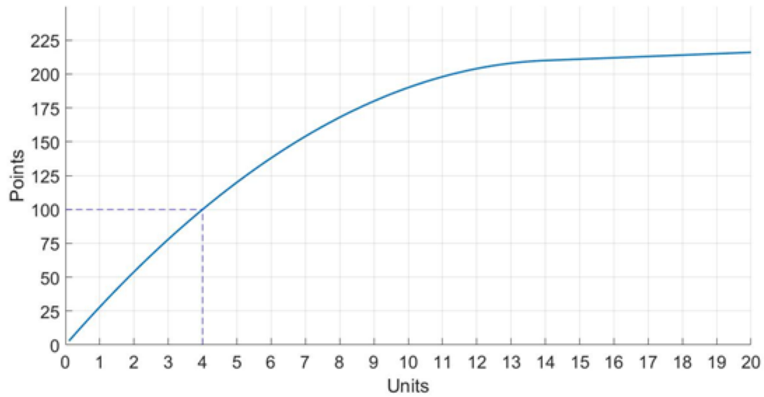
You may revise your choices at any moment before the round ends. During a round, you will also be informed about every other participant's most recent decision (units bought and formed links), which will be updated every 5 seconds or whenever you change your own choice.

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 $\times$ 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 $\times$ 1 from P1 + 0.5 $\times$ 1 from P2 + 0.5 $\times$ 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 $\times$ 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 $\times$ 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.

D Network game interface

The decision making interface used in the experiment is similar across all treatments. More specifically, Figure A3 illustrates a (fictitious) example of a subject's computer screen in Treatment **Baseline100**. 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), the subject's own effort, which can be modified via the slider, and a comprehensive description of the subject's own payoff. Information about payoffs include gross earnings (output of function $f(.)$), the cost of effort (own effort multiplied by c), the cost of linking (number of links 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, 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).

Similarly, Figure A4 illustrates a (fictitious) example of a subject's computer screen in Treatment **PayInfo100**. The only difference with the decision screen from Figure A3 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.

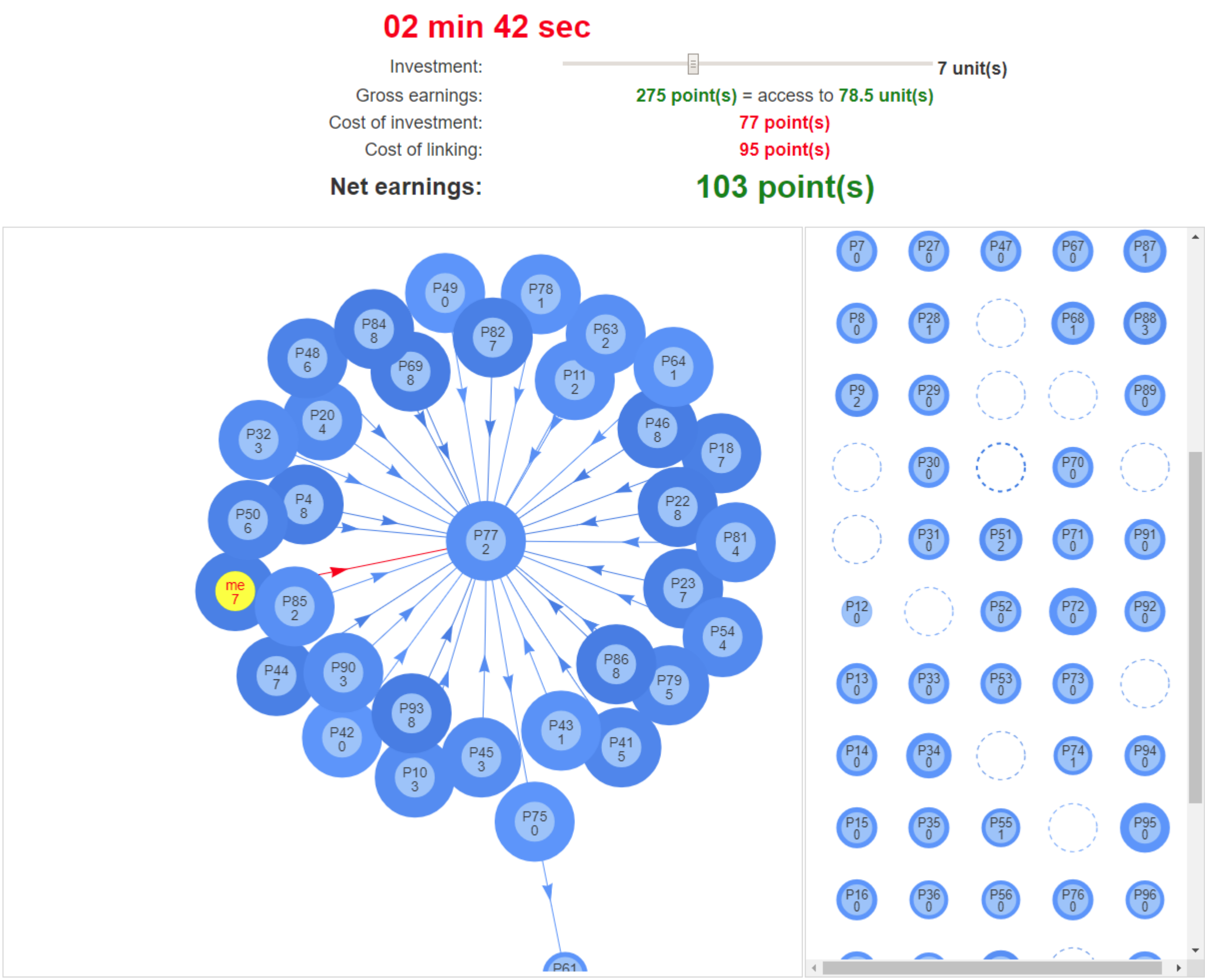


Figure A3: Example of decision screen for Treatment **Baseline100**

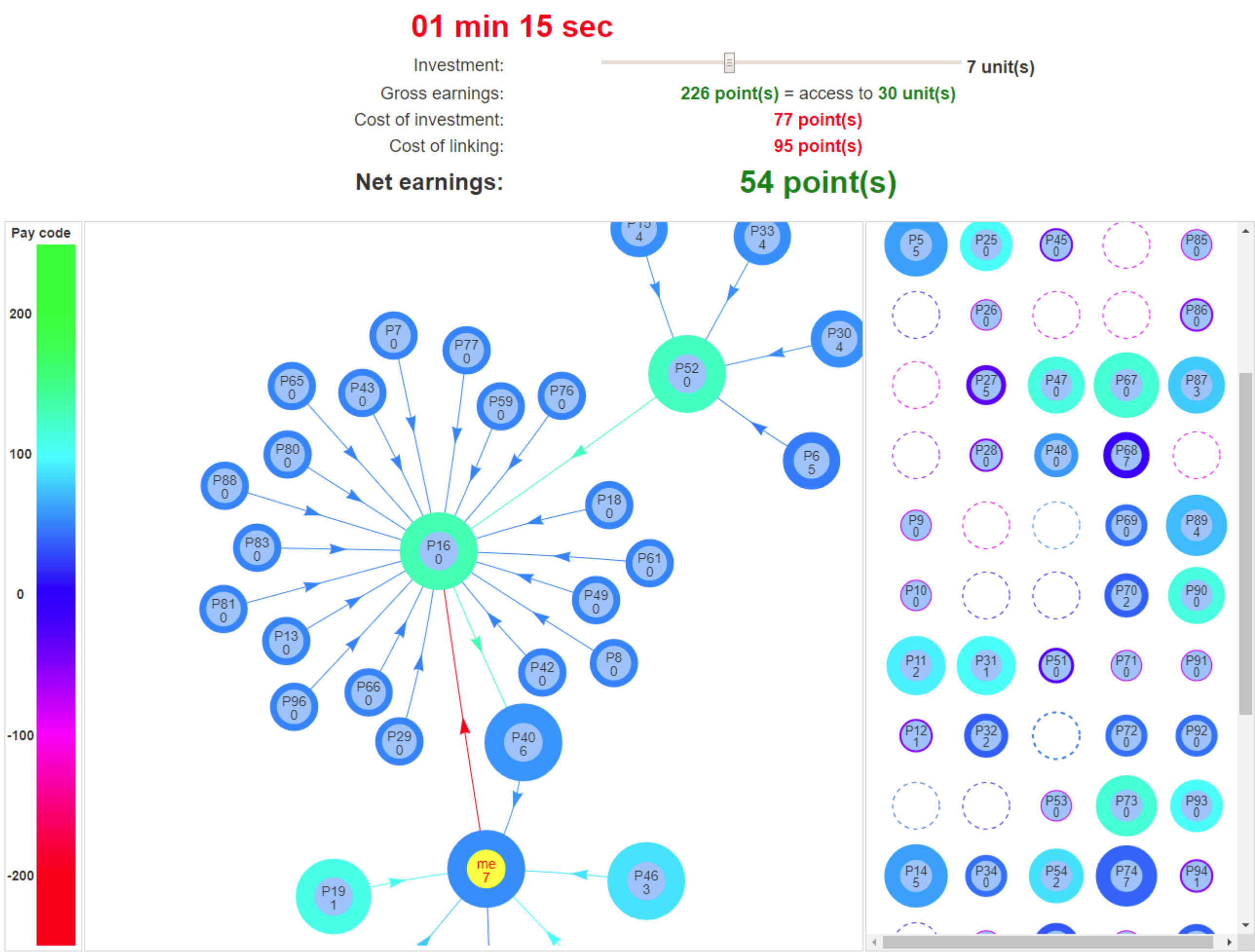


Figure A4: Example of decision screen for Treatment **PayInfo100**

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.

A Comprehension check

In order to assess the subjects' comprehension of the network game played during the experiment, we provided 5 questions, each of which with a unique correct answer. Each correct answer was rewarded with 0.1 euro for the subject.

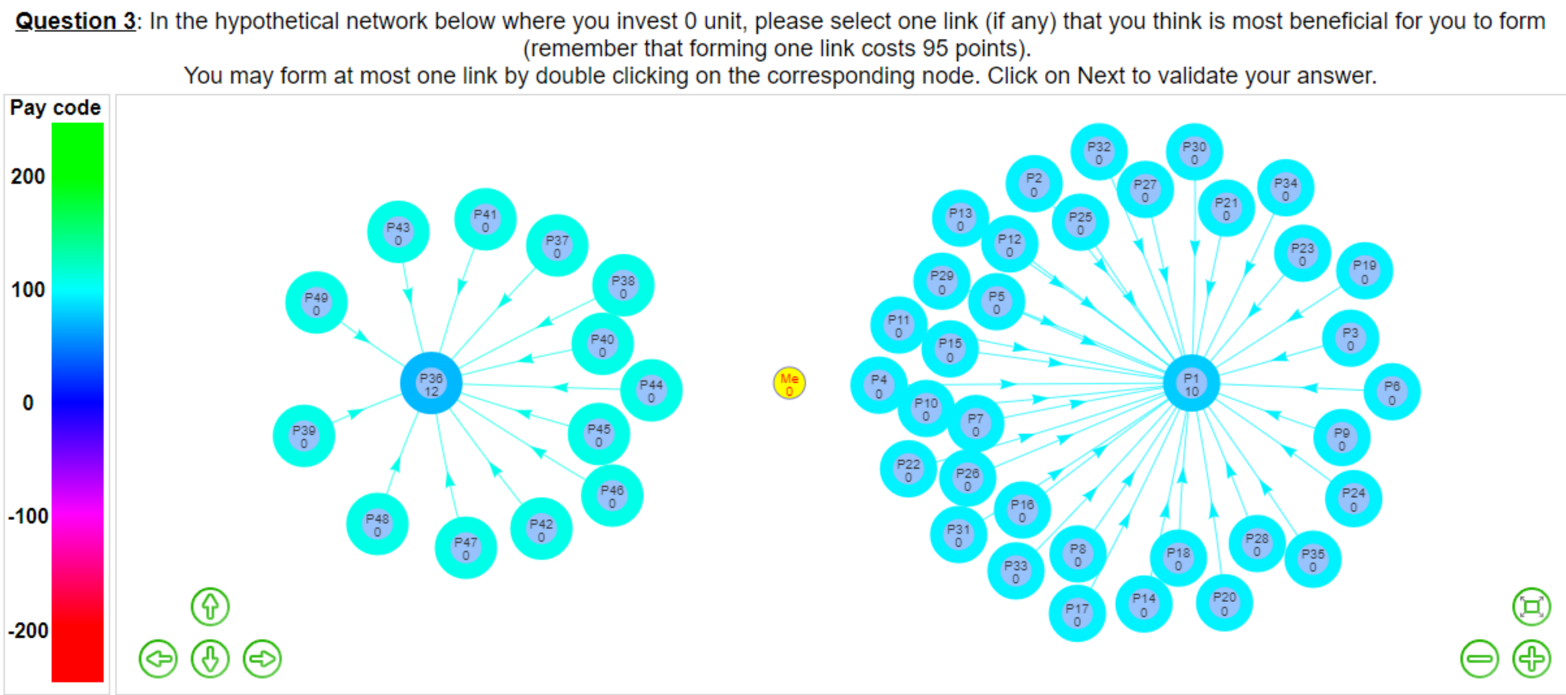
The following first 2 questions were used across all treatments (correct answers are "11 pts" to question 1, and "95 pts" to question 2).

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

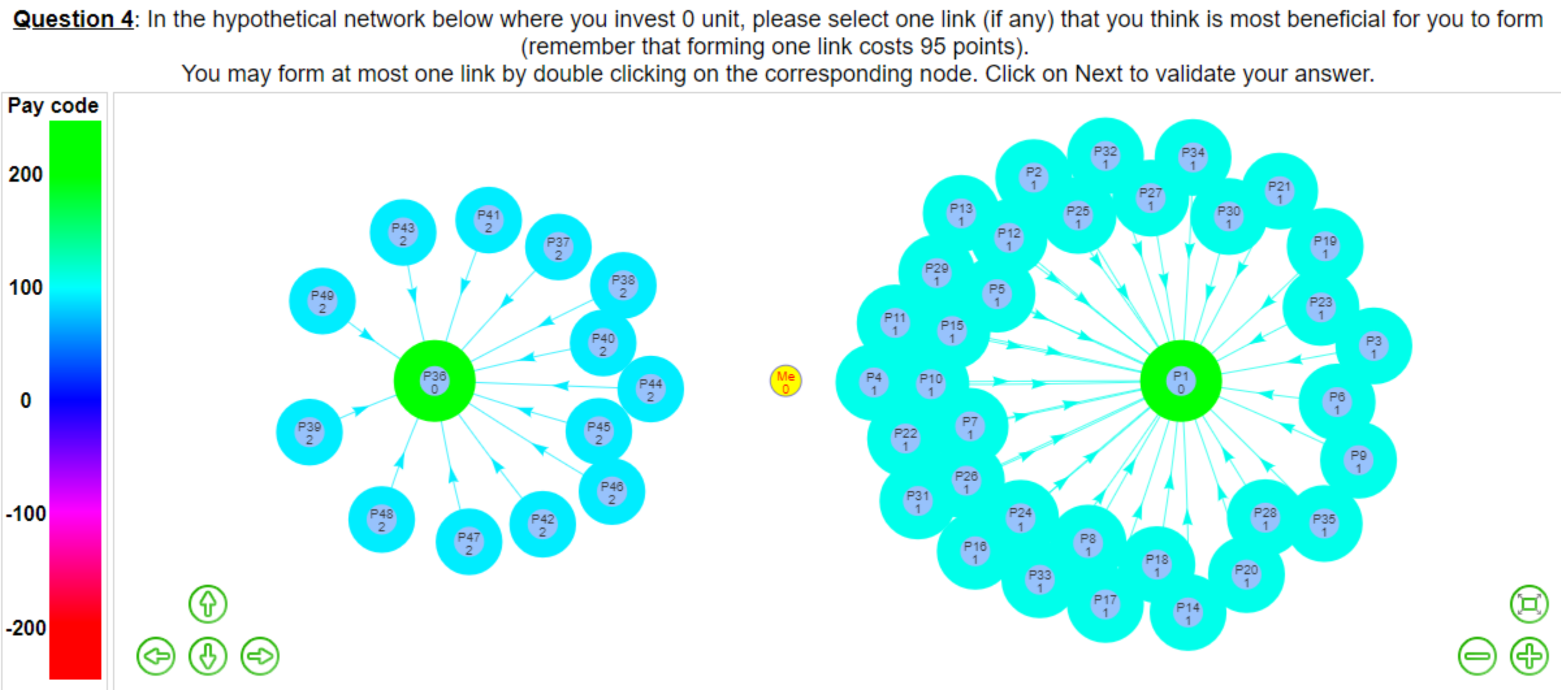
Question 2: In the previous game, how many points did forming a link cost you?
- 0 pt
- 25 pts
- 45 pts
- 65 pts
- 95 pts
- 115 pts
- more than 115 pts

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

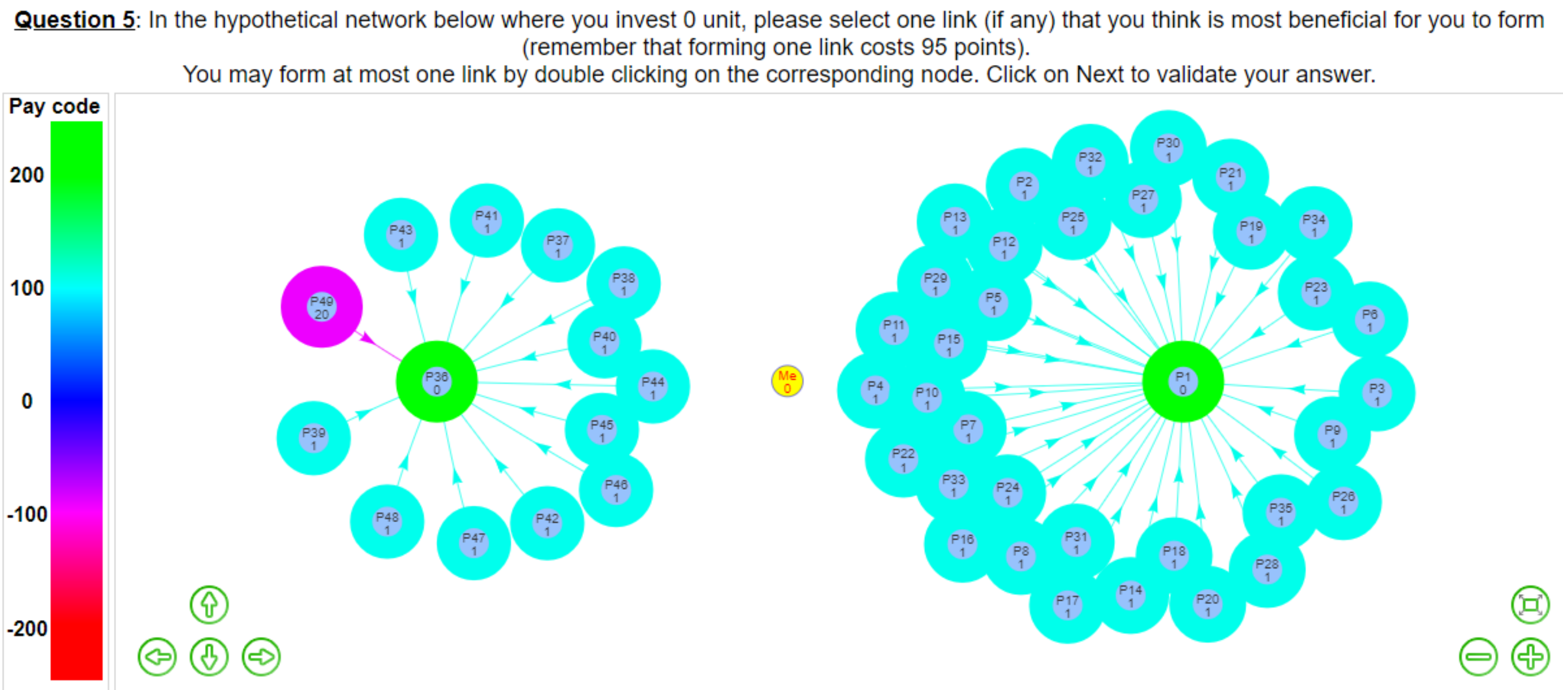


however adapted only in other treatments where $n > 4$ by again matching the number of nodes to the group size in the experiment. The reason for filtering the small group treatments (with $n = 4$) is that the limited number of nodes did not allow representing the corresponding scenarios. As before, these questions were also adapted to the baseline treatments by simply removing the colors to match the design of the actual game that subjects played.



B 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:

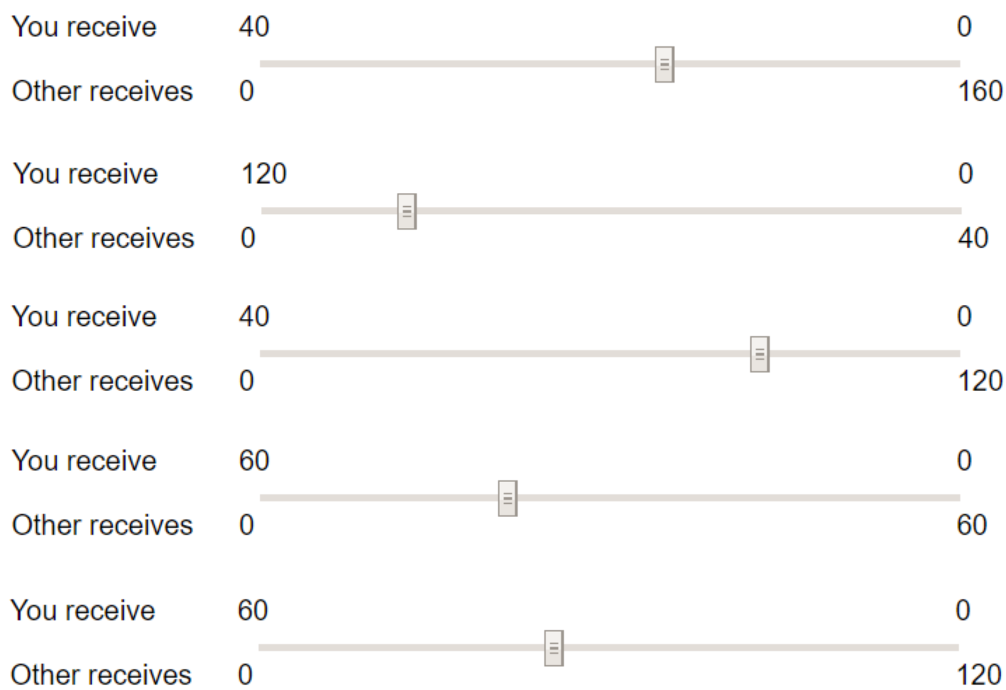


Note however that each question was presented in a different screen, and the order of presentation was randomized for every subject. Furthermore, 50 points were worth 1 euro both the subject, and the other anonymous external participant. Detailed instructions provided to the subjects, as well as a screenshot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

C 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.



D 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.

Question 1

Please select your preferred allocation on the slider below
(values are in points, with 50 points = 1 euro):

You receive	17
Other receives	93

You receive	40	0
Other receives	0	160

Next

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 Additional tables and figures

A Regression tables

Table A1: Scale effects on effort and payoff in the baseline treatments

	Mean effort			Median payoff		
	most connected	2nd most connected	others	most connected	2nd most connected	others
$N = 50$	6.61***	7.27***	0.32	-40.81***	-51.09**	28.82***
	(1.08)	(1.41)	(0.32)	(10.20)	(23.61)	(1.73)
Average in small group	8.77	5.24	2.65	86.50	81.00	85.00
Number of observations	60	60	1120	60	60	1120
R-squared	0.61	0.59	0.04	0.39	0.23	0.11

Notes: Robust standard errors (clustered by individual subject in the regression analysis) are reported in parenthesis. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively. All regressions include a constant, round dummies, age, female, education, comprehension test score, experimental measures of risk aversion and altruism, and Big 5 personality.

Table A2: Scale effects on effort and payoff in the baseline treatments

	Mean effort			Median payoff		
	most connected	2nd most connected	others	most connected	2nd most connected	others
$N = 100$	6.64***	11.06***	0.88***	16.54	-25.41*	53.20***
	(1.54)	(1.10)	(0.32)	(29.95)	(14.54)	(2.77)
Average in small group	8.77	5.24	2.65	86.50	81.00	85.00
Number of observations	55	55	1630	55	55	1630
R-squared	0.62	0.83	0.04	0.20	0.38	0.14

Notes: Robust standard errors (clustered by individual subject in the regression analysis) are reported in parenthesis. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively. All regressions include a constant, round dummies, age, female, education, comprehension test score, experimental measures of risk aversion and altruism, and Big 5 personality.

Table A3: Scale effects on outdegree in the baseline treatments

	Mean outdegree		
	most connected	2nd most connected	others
Large group	1.03***	0.75*	0.24***
	(0.35)	(0.40)	(0.05)
Average in small group	0.20	0.62	0.90
Number of observations	75	75	2590
R-squared	0.38	0.41	0.03

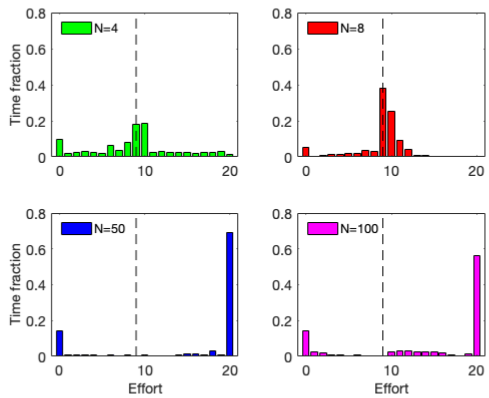
Notes: Robust standard errors, clustered by individual subject, are reported in parenthesis. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively. All regressions include a constant, round dummies, age, female, education, comprehension test score, experimental measures of risk aversion and altruism, and Big 5 personality.

Table A4: Treatment effects on effort and payoffs

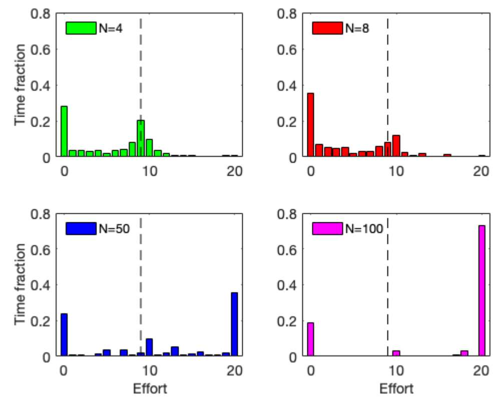
	Mean effort			Median payoff		
	most connected	2nd most connected	others	most connected	2nd most connected	others
Payoff info	-0.75	0.52	0.00	6.71	-12.75***	-10.56***
	(0.77)	(0.70)	(0.36)	(11.54)	(4.53)	(1.97)
Large group	6.30***	8.41***	0.62**	-30.33*	-42.76**	36.20***
	(1.04)	(1.19)	(0.30)	(17.08)	(17.34)	(1.90)
Payoff info × Large group	-9.24***	-9.00***	-0.91**	119.24***	120.76***	-14.07***
	(1.41)	(1.63)	(0.39)	(29.18)	(29.02)	(2.30)
Mean or median in large group baseline	15.12	13.22	3.22	59.00	47.00	126.50
Number of observations	150	150	5180	150	150	5180
R-squared	0.53	0.51	0.05	0.09	0.17	0.09

Notes: Robust standard errors (clustered by individual subject in the regression analysis of efforts) are reported in parenthesis. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively. All regressions include a constant, round dummies, age, female, education, comprehension test score, experimental measures of risk aversion and altruism, and Big 5 personality.

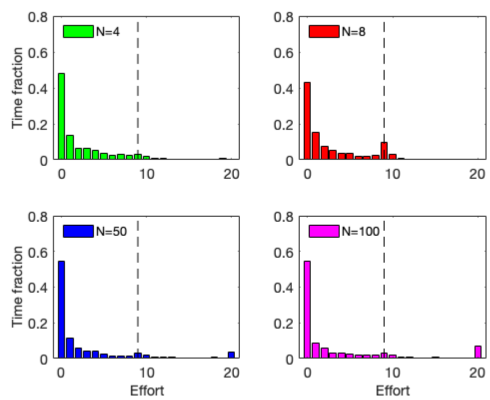
B Additional figures



(a) the 1st most connected

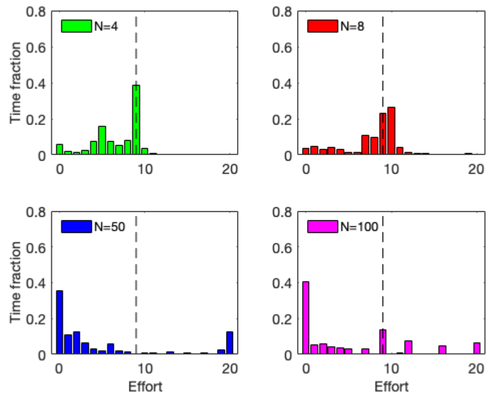


(b) the 2nd most connected

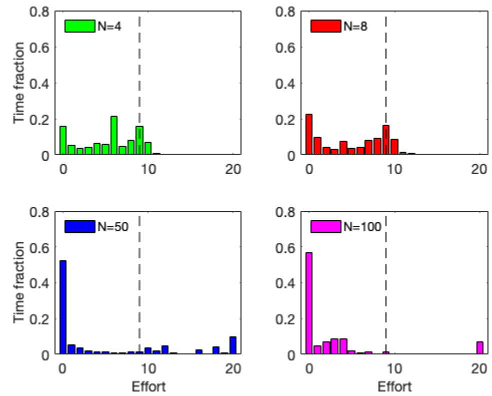


(c) the others

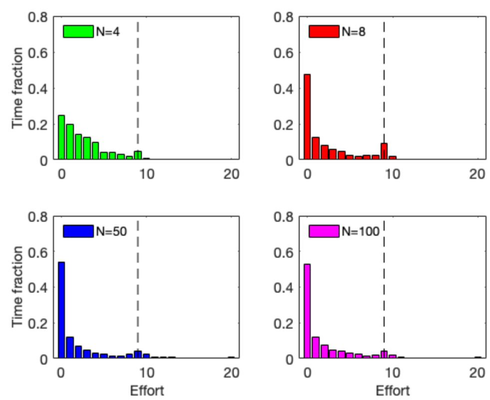
Figure A5: Distribution of efforts in the baseline treatment



(a) the 1st most connected



(b) the 2nd most connected



(c) the others

Figure A6: Distribution of efforts in the payoff information treatment

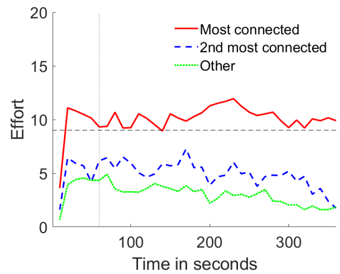
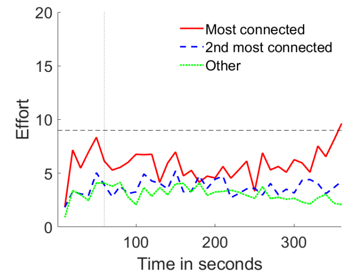
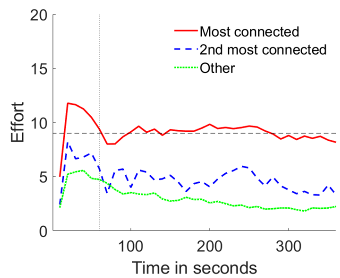
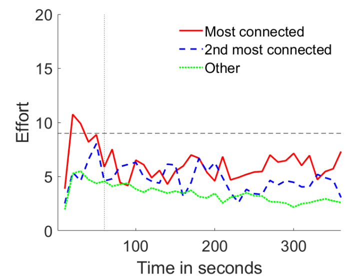
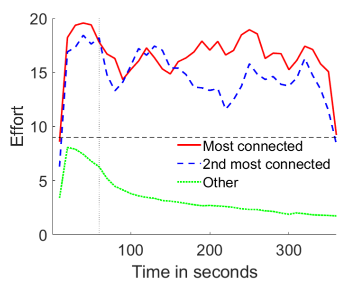
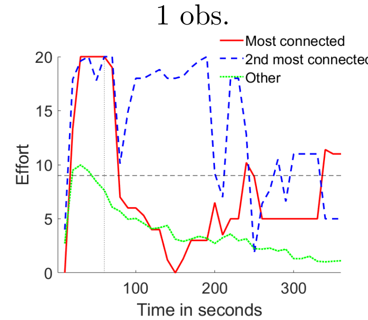
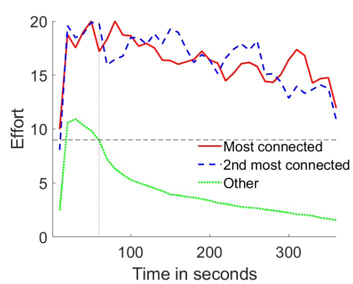
<table>
<tr><th rowspan="2">N</th><th colspan="3">Types of Outcomes</th></tr>
<tr><th>Pure Influencer</th><th>Pure Connecteor</th><th>Other</th></tr>
<tr><td>4</td><td>14 obs.
</td><td>NA</td><td>6 obs.
</td></tr>
<tr><td>8</td><td>16 obs.
</td><td>NA</td><td>4 obs.
</td></tr>
<tr><td>50</td><td>19 obs.
</td><td>NA</td><td>1 obs.
</td></tr>
<tr><td>100</td><td>15 obs.
</td><td>NA</td><td>NA</td></tr>
</table>

Table A5: Time series of efforts for the two different types of subjects in different types of outcomes (Baseline treatments)

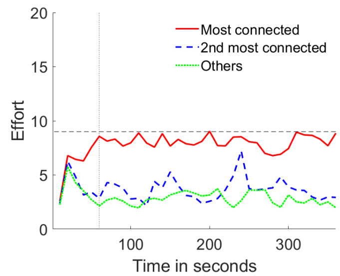
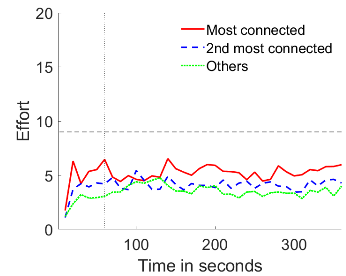
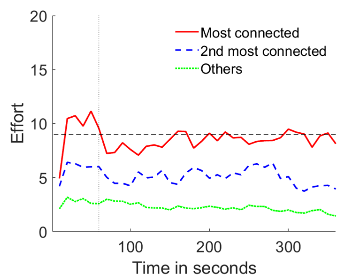
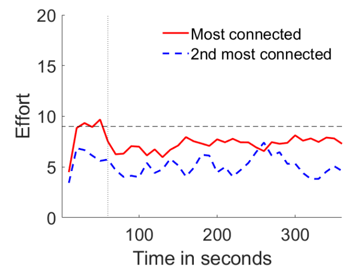
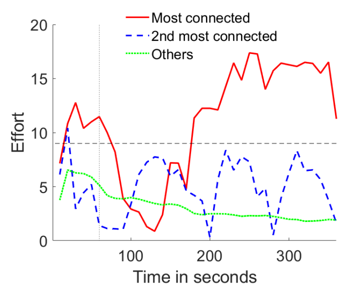
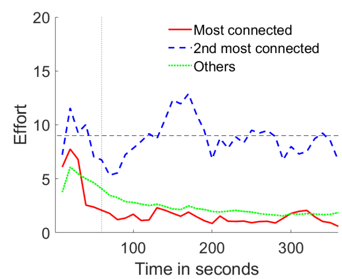
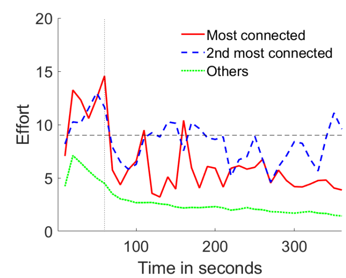
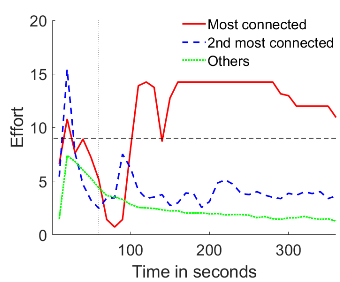
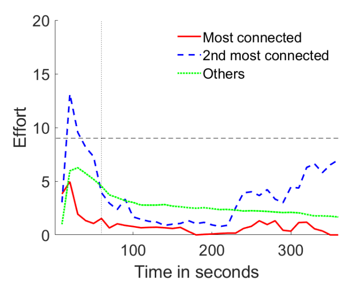
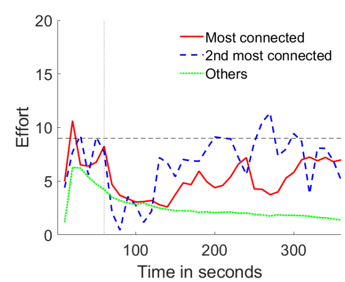
<table>
<tr><th></th><th colspan="3">Types of Outcomes</th></tr>
<tr><th>N</th><th>Pure Influencer</th><th>Pure Connecteor</th><th>Other</th></tr>
<tr><td>4</td><td>7 obs.
</td><td>NA</td><td>13 obs.
</td></tr>
<tr><td>8</td><td>16 obs.
</td><td>NA</td><td>4 obs.
</td></tr>
<tr><td>50</td><td>4 obs.
</td><td>9 obs.
</td><td>7 obs.
</td></tr>
<tr><td>100</td><td>4 obs.
</td><td>6 obs.
</td><td>5 obs.
</td></tr>
</table>

Table A6: Time series of efforts for the two different types of subjects in different types of outcomes (PayInfo treatments)

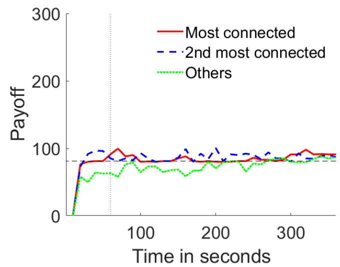
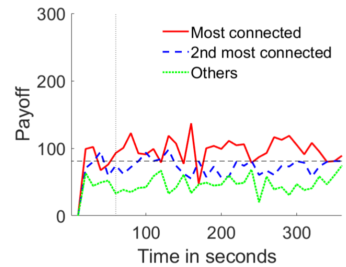
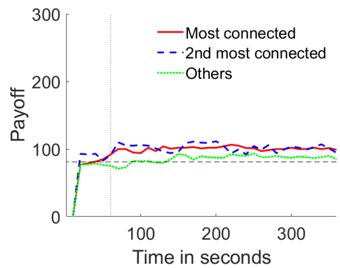
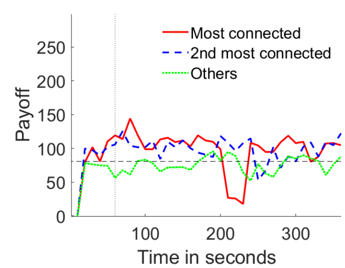
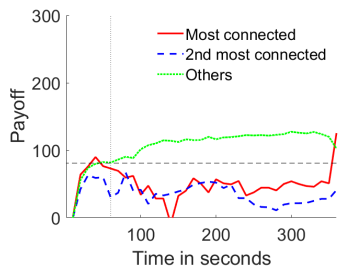
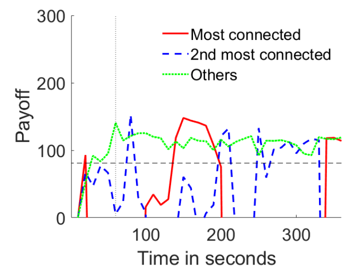
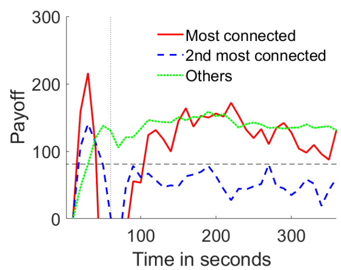
<table>
<tr><th></th><th colspan="3">Types of Outcomes</th></tr>
<tr><th>N</th><th>Pure Influencer</th><th>Pure Connecteor</th><th>Other</th></tr>
<tr><td>4</td><td>14 obs.
</td><td>NA</td><td>6 obs.
</td></tr>
<tr><td>8</td><td>16 obs.
</td><td>NA</td><td>4 obs.
</td></tr>
<tr><td>50</td><td>19 obs.
</td><td>NA</td><td>1 obs.
</td></tr>
<tr><td>100</td><td>15 obs.
</td><td>NA</td><td>NA</td></tr>
</table>

Table A7: Time series of median payoffs for the two different types of subjects in different types of outcomes (Baseline treatments)

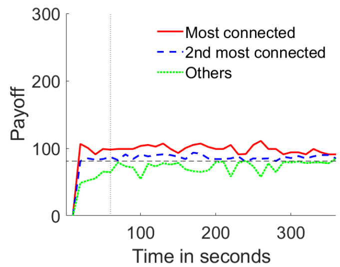
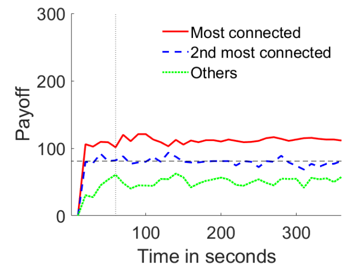
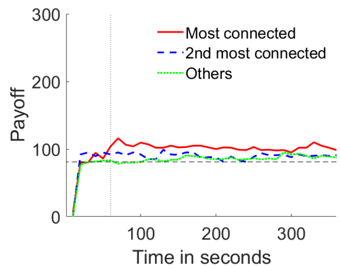
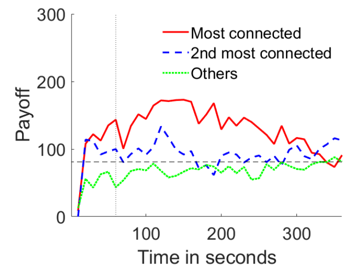
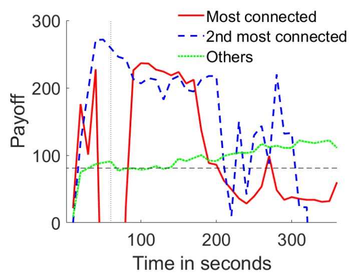
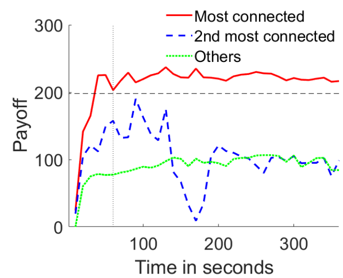
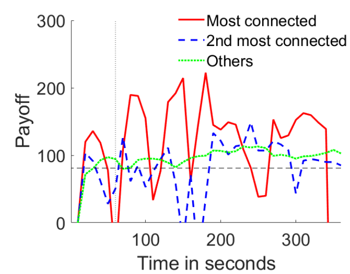
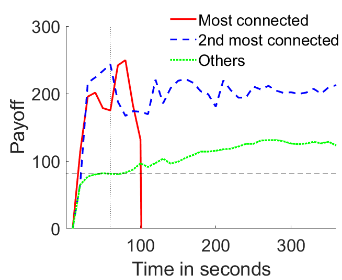
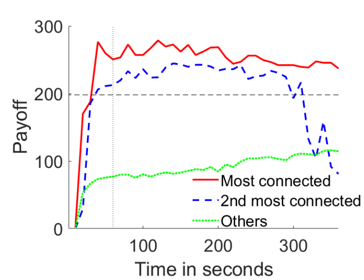
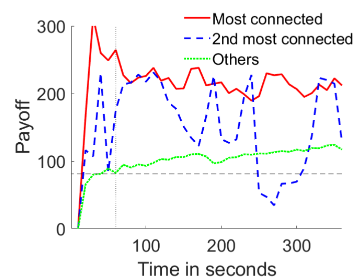
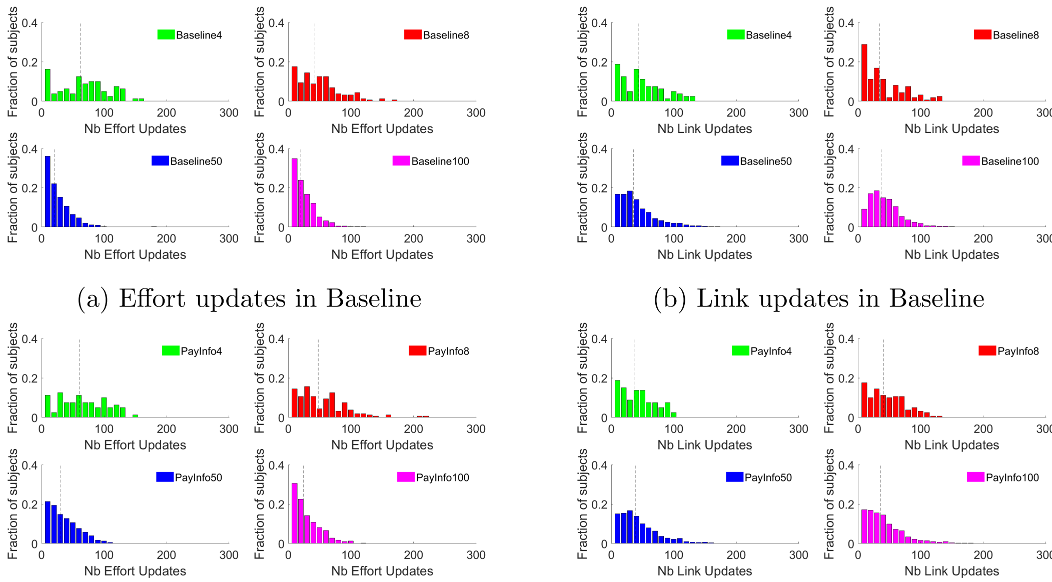
N	Types of Outcomes Pure Influencer	 Pure Connecteor	 Other
4	7 obs. 	NA	13 obs. 
8	16 obs. 	NA	4 obs. 
50	4 obs. 	9 obs. 	7 obs. 
100	4 obs. 	6 obs. 	5 obs. 

Table A8: Time series of median payoffs for the two different types of subjects in different types of outcomes (PayInfo treatments)



(a) Effort updates in Baseline

(b) Link updates in Baseline

(c) Effort updates in PayInfo

(d) Link updates in PayInfo

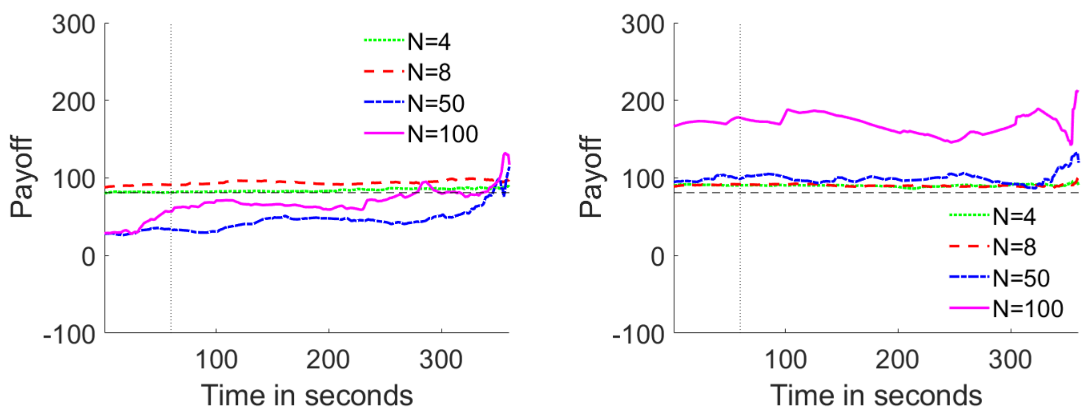
Figure A7: Distribution of choice updates

		Outcome Types								
		Pure Influencer			Pure Connector			Other		
Treatment	N	#Obs	Hub	Others	#Obs	Hub	Others	#Obs	Hub	Others
Baseline	4	14	10.2	3.6	0	.	.	6	5.8	3.3
	8	16	9	2.9	0	.	.	3	5.9	3.4
	50	14	15.9	2.8	0	.	.	0	.	.
	100	11	15.5	3.4	0	.	.	0	.	.
PayInfo	4	7	8.1	3.1	0	.	.	13	5.25	3.8
	8	16	8.4	2.6	0	.	.	4	5.3	3.5
	50	4	11.1	2.8	9	1.4	2.1	4	6	2.2
	100	0	.	.	6	0.6	2.4	4	4.5	2.3

Table A9: Mean effort in last 5 minutes across different types of Outcomes (excluding outliers, i.e., rounds where the most connected individual earns less than -100 points during at least 60 seconds in the last 5 minutes)

		Nb of Rounds				
Treatment	N	1	2	3	4	5
Baseline	4	0	12.5	31.2	50	
	8	18.7	25	21.9	12.5	18.7
	50	13	3.5	1.5	1	0.5
	100	5.7	1	0.3	0.3	0.3
PayInfo	4	0	0	0	0	1
	8	3.1	28.1	25	15.6	21.9
	50	21	4.5	1.5	0.5	0
	100	9.3	0.3	0.7	0	0

Table A10: Fraction of subjects (%) becoming the most connected individual for at least 10 seconds in any number of rounds, across treatments



(a) Baseline (b) PayInfo

Figure A8: Continuation payoffs