# Efficiency and equilibrium in network games: An experiment

Edoardo Gallo*and Chang Yan†‡

February 2015

**Abstract**

The tension between efficiency and equilibrium is a central feature of economic systems. In many contexts, social networks mediate this trade-off: an individual's network position determines equilibrium play, and social relations allow coordination on an efficient norm. We examine this trade-off in a network game with a unique Nash equilibrium, but such that agents can achieve a higher payoff by following a "collaborative norm". Subjects establish and maintain a collaborative norm in the circle, but the norm weakens with the introduction of one asymmetric node in the wheel. In complex and asymmetric networks of 15 and 21 nodes, the norm disappears and subjects' play converges to Nash on every node. We provide evidence that subjects base their decisions on their degree, rather than the overall network structure. Methodologically, the paper shows the capabilities of *UbiquityLab*: a novel platform to conduct interactive experiments online with a large number of participants.

**Keywords:** network, online experiment, network game, strategic complements.
**JEL:** C99, D03, D85, Z13.

*University of Cambridge and Queens' College, Cambridge CB3 9ET, UK. Email: *edo@econ.cam.ac.uk*
†University of Oxford, Oxford OX1 3QD, UK. Email: *chang.yan@cs.ox.ac.uk*

‡We thank Marco Felici and Joyce Ong for outstanding research assistance. We are grateful to Stefano Caria, Syngjoo Choi, Vessela Daskalova, Sanjeev Goyal, Margaret Meyer and Luca Stanca for helpful comments and suggestions. Thanks to seminar participants at the University of Oxford, Cambridge-INET, University of Cambridge, Università degli Studi di Milano-Bicocca, Seoul National University, UC Irvine, University of Essex, and the Psychometrics Center (Cambridge); and to conference participants at the Social Networks and Matching Processes Workshop (CSAE, Oxford), the Learning in Social Networks Workshop (IESE, Barcelona), the Keynes Fund Research Day (Cambridge), and the 2013 ESI Autumn Workshop (Max Planck, Jena). Generous research support by the Keynes Fund for Applied Economics in Cambridge, the Balliol Interdisciplinary Institute and the Oxford-Man Institute of Quantitative Finance is thankfully acknowledged.

# 1 Introduction

The tension between efficiency and equilibrium is a central feature of social and economic systems. In some contexts, efficiency can be achieved and sustained through the establishment of a norm: a prominent example is the provision of a public good where we have abundant empirical (Ostrom [1990]) and experimental (Ledyard [1995]) evidence that individuals manage to coordinate away from the zero contribution equilibrium. Norms are intertwined with a society's culture and social structure as their sustenance depends on a set of common beliefs and expectations which are often maintained through social relations[1]. Social structure can also be a key determinant of an equilibrium: recent theoretical research has shown that the position of an individual in a social network is a determinant of equilibrium behavior in many economic contexts[2].

Despite the importance of social structure for equilibrium behavior and the establishment of norms, we have scarce evidence of how conducive different social structures are to the establishment and sustenance of a norm in contexts where position in the social structure determines equilibrium play. A primary reason for this gap is the largely unresolved challenge of unambiguously identifying the causal link from position in the social structure to behavior: social econometrics is still in its infancy and it struggles to validate the theory using empirical data. Experiments are a natural tool to pin down the causal link, but most of the experimental research on networks in economics focuses on networks that are small in size and/or have a stylized structure, which limits the testing of the rich relation between network position and behavior predicted by the theory.

This paper addresses this gap by investigating experimentally a game of strategic complements played on sizable networks with a non-trivial structure. We have chosen this specific game for three reasons. First, it has a unique Nash equilibrium in which agents' play is determined by their position in the network: they exert efforts that are proportional to their Bonacich centrality, a metric that captures the number of connections of an agent, her neighbors, her neighbors' neighbors, and so on. Second, in all the networks we examine, it is possible for agents to achieve payoffs higher than Nash play if everyone coordinates on exerting higher efforts, which we dub a "collaborative norm". Third, this type of game captures a broad range of phenomena involving collaboration among members of a group with complementary skills such as teamwork in organizations, criminal activity, and coauthorships in academia.

The experiment consists of playing 40 rounds of the game on a given network structure. In a round each subject is randomly assigned to a different node of the network, and picks a level of effort, which can be any integer in the $[0, 100]$ range. There are 4 treatments that differ in the network the subjects are exogenously assigned to. Figure 1 illustrates the 4 networks: a circle network $\mathbf{g}^{\bigcirc}$ and a wheel network $\mathbf{g}^{\otimes}$ with 9 nodes each, a 15-node

[1]See Young [2014] for a comprehensive review of research on social norms in economics.

[2]See Goyal [2007] and Jackson [2008] for comprehensive reviews.

network $\mathbf{g}^{15}$, and a 21-node network $\mathbf{g}^{21}$. We conduct the experiment using *UbiquityLab*: a platform we built to conduct interactive experiments online with a large number of participants.

Our first finding is that in simple networks subjects are able to establish and maintain a collaborative norm where they exert efforts higher than equilibrium play. In the simple and symmetric $\mathbf{g}^{\bigcirc}$ network, subjects converge to average efforts that are in between the effort predicted by Nash play and the most efficient collaborative norm for the group. As a minimal degree of asymmetry is introduced in the $\mathbf{g}^{\otimes}$ network, the collaborative norm survives but it is weakened: efforts are closer to Nash play than to the most efficient collaborative norm. Finally, in the more asymmetric and complex[3] $\mathbf{g}^{15}$ and $\mathbf{g}^{21}$ networks, subjects are unable to sustain any kind of collaborative norm.

The second result is that subjects converge to the Nash equilibrium in all positions of the $\mathbf{g}^{15}$ and $\mathbf{g}^{21}$ networks. In spite of the presence of highly profitable collaborative norms, it is likely that the asymmetry and complexity of the networks make it challenging for subjects to achieve coordination. It is somewhat remarkable that subjects are able to learn to play the equilibrium given the size of the strategy space and the number of different positions in the network structure. For instance, the $\mathbf{g}^{21}$ network has 5 types of positions and the Nash equilibrium predictions range from 33 to 78. In spite of this complexity, subjects' average efforts on all node types in the $\mathbf{g}^{15}$ and $\mathbf{g}^{21}$ networks are statistically indistinguishable from the Nash predictions.

These findings show that position in the network structure is highly predictive of subjects' decisions in networks of significant complexity. The full validation of the Nash predictions in these complex networks is somewhat surprising, given the intricate relation between Nash play and position with respect to the overall network structure. It is clear from the data that subjects do not start by playing Nash, so a natural question is how they are able to converge to Nash over time. We investigate the hypothesis that subjects simplify the task by focusing on the local features of the network structure. A natural metric to capture the position in the network using local information is the number of connections, or degree, which is highly correlated to Bonacich centrality in most networks.

Our third finding is that subjects base their decisions on the degree: they exert higher effort the larger is the degree of the node they are assigned to. We specifically designed the $\mathbf{g}^{21}$ network to be able to separate play according to degree from Bonacich. Equilibrium play in network $\mathbf{g}^{21}$ is such that subjects in node type $P$ (degree 1) exert the same effort as subjects in the central node $F$ (degree 2), and a lower effort than subjects assigned to node type $C$ (degree 2) that is a neighbor of $F$. In the data we observe that subjects' play in node types $P$ and $C$ is indistinguishable from Nash, but subjects in the central node $F$ choose an effort that is indistinguishable from the effort in node type $C$ and qualitatively

[3]Throughout this paper we use the terms "complex" and "complexity" in a non-technical fashion to convey the intuitive fact that the structure of networks $\mathbf{g}^{15}$ and $\mathbf{g}^{21}$ has more asymmetries than networks $\mathbf{g}^{\bigcirc}$ and $\mathbf{g}^{\otimes}$.

higher than the effort in node type $P$, as expected if they based their decision on degree. Moreover, in a panel regression analysis, if we restrict the attention to nodes $P$, $C$ and $F$, then only degree remains a significant predictor of subjects' choices. Thus, in this context, it suffices to have information about a simple network metric, the degree of an individual, to predict behavior, rather than having to collect information about the whole network structure.

These results highlight the importance of going beyond small and stylized networks to investigate how network structure determines behavior. As we will see below, experimental research on networks in economics has so far failed to investigate behavior in a game with a rich strategy space played on large networks with a complex structure. This simplification comes at the cost of largely failing to test the behavioral validity of the relation between a rich network structure and behavior predicted by the theory. We find that in stylized networks with a few nodes and/or asymmetries, such as the circle and the wheel, subjects are able to process the network structure and coordinate on norms of behavior. However, in complex networks with many asymmetries, subjects find it more challenging to achieve the coordination necessary to establish and maintain a norm, and they mainly base their decisions on the local network structure to reduce the cognitive cost of the task of processing complex network information. It is crucial to investigate experimentally the generalizability of these findings to different contexts if we want to give novel input for theoretical work and enhance its behavioral validity.

The remaining part of this section surveys the related literature. Section 2 presents an overview of the theory. Section 3 gives the details of the experimental design and lists our hypotheses. Section 4 analyzes the experimental data and tests the validity of the hypotheses. Section 5 conducts some robustness checks of our results, and section 6 concludes. The paper also includes three appendices: appendix A provides further details on the web platform we developed to conduct the experiment, appendix B presents some further analysis of the data, and appendix C contains the screenshots of the instructions for the experiment and the post-experimental questionnaire.

## 1.1 Literature review

The study of whether individuals converge to equilibrium play in various games has been a major theme of inquiry since the beginning of experimental economics. In many contexts, a force that counters convergence to equilibrium play is the emergence of norms that lead to more efficient outcomes. Despite the large body of work along these themes, only recently researchers have turned their attention to how the overall structure of social networks and the position of individuals in the network plays a role in mediating the tension between equilibrium play and efficiency. Our paper is a contribution to this more recent literature and it ties the role of social structure to some familiar findings in the experimental literature.

First of all, this paper contributes to examining the impact of asymmetries among

individuals on the establishment and sustenance of a social norm. In the extensive experimental literature on public good games, several researchers have investigated the effect of introducing heterogeneity among subjects on the positive contribution norms observed in the benchmark case with homogeneous subjects. Ledyard [1995] surveys several papers and finds that the introduction of heterogeneity in preferences or endowments decreases the level of contributions. For instance, Rapoport and Suleiman [1993] find that groups where individuals have different endowments are less successful in providing the public good compared to groups where everyone has the same endowment. To our knowledge, our paper is the first one to investigate position in the network structure as a source of asymmetry among agents in the context of the establishment and sustenance of a social norm. Our results on the gradual breaking down of the norm as more complexity, or asymmetry, is added to the network is consistent with results in the literature, albeit in the context of a different game where the role of norms has received little attention before.

Our paper is the first comprehensive experimental investigation of the game of strategic complements on a network in Ballester et al. [2006]. The lack of experimental work on this topic is surprising because the causal relation between network position and behavior in this game is well-understood at theoretical level since Ballester et al. [2006], which is one of the seminal papers in the network literature in economics. Moreover, this type of game captures a wide range of applications in which social networks matter for behavior, and that have been extensively studied in sociology and other disciplines. Charness et al. [2014] is the most closely related paper in this vein of work. They investigate network games of strategic complements and substitutes, both with complete and partial knowledge of the network, in which the action space is binary. The focus of their paper is a test of the comparative statics predictions in the Galeotti et al. [2010] paper which introduces incomplete information in network games. Moreover, their choice of binary actions means that all outcomes are corner solutions, which makes it challenging to investigate the rich relation between equilibrium play and network position in the Ballester et al. [2006] framework.

More broadly, this paper contributes to a growing literature in experimental economics investigating games played on exogenously given networks. Selected contributions include bargaining (Charness et al. [2007], Gallo [2014]), trading (Gale and Kariv [2009], Choi et al. [2014]), public goods (Choi et al. [2011], Rosenkranz and Weitzel [2012]), prisoner's dilemma (Cassar [2007], Gallo and Yan [2014]), and coordination (Cassar [2007]) games. A common characteristic of these papers is that they largely focus on small and stylized networks that do not allow an in-depth investigation of the intricate causal relation from network position to behavior[4]. In contrast, there are three elements in our set-up whose combination allows the exploration of this relation: the size of the networks, the presence

[4]Outside of economics there is an experimental literature that investigates games played on large networks (e.g. Gracia-Lázaro et al. [2012], Kearns et al. [2006], Kearns et al. [2009]), but the focus is on the effect of the overall network structure rather than the relation between position in the network and behavior.

of asymmetry across nodes, and the large strategy space. To our knowledge, the presence of all these three elements is a distinctive feature of our paper. Charness et al. [2014] has three treatments with asymmetric networks of similar size to ours, but the strategy space is binary. Choi et al. [2014] explore large networks and subjects can choose from a large strategy space, but the networks are highly symmetric. Moreover, our findings show that the analysis of behavior on both stylized and complex networks leads to novel insights compared to an investigation limited to either type.

A recent literature has begun to examine how individuals process information about complex networks and how they use this information to take decisions. Social psychologists have used a survey-based methodology proposed by Krackhardt [1987] to map individuals' cognitive perception of their social environment and they have identified a variety of cognitive biases (see, e.g., Kumbasar et al. [1994], Casciaro [1998]). In a controlled experimental setting, Dessi et al. [2014] show that individuals utilize heuristics to memorize and recall information about networks of 15 individuals, and this leads to systematic biases. Despite the evidence that individuals resort to heuristics to cope with the complexity of network information, there are still very few studies that investigate how this affects individuals' decisions by analyzing treatments that differ in the degree of network information provided to participants. Two exceptions are Charness et al. [2014], who show that incomplete information about the network structure makes it more difficult to coordinate on an efficient equilibrium, and Gallo and Yan [2014] who show that individuals use information about the whole network to form separate communities in the context of a prisoner's dilemma game with endogenous network formation. In our paper, we do not have treatments with different network information, but we provide evidence that individuals cope with the complexity of the network by basing their decisions on local network information rather than the overall network structure.

In summary, this paper makes 4 main contributions. First, it sheds light on the relation between network structure and the establishment of a social norm that leads to a better aggregate outcome than equilibrium play: the norm emerges and survives in stylized and symmetric networks, but it disappears in more complex networks that generate asymmetries across subjects. Second, this is the first experimental test of the seminal result in Ballester et al. [2006] for games of strategic complements on a network: the theory turns out to be highly predictive of subjects' choices in complex networks. Third, it increases our understanding of how individuals cope with the complexity of a social structure that matters in their decisions: here they mainly focus on the local rather than the global structure of the network, which explains the single deviation from the Nash predictions that we observe in one of the two treatments with a complex network. Finally, it shows the potential of UbiquityLab and the novel methodology of interactive online experiments.

## 2 Theory

The theory underlying this experiment is well-established since the seminal article by Ballester et al. [2006]. In this section we provide a brief overview that serves the purposes of introducing relevant notation, providing intuition for the key result, and briefly mentioning some recent developments.

Consider a set of $N = \{1, ..., n\}$ agents who are the nodes of a network $\mathbf{g}$. The network is undirected and unweighted[5] so for any $i, j \in N$ we have that either there is a link $g_{ij} = g_{ji} = 1$, or they are not connected, i.e. $g_{ij} = g_{ji} = 0$. Each agent $i$ selects an effort level $e_i$ and obtains utility:

$$u_i(e_1, ..., e_n) = \alpha e_i - \frac{\beta}{2} e_i^2 + \lambda \sum_{j=1}^{n} g_{ij} e_i e_j \tag{1}$$

where $\alpha > 0$ and $\beta > 0$ capture the strength of the idiosyncratic benefits and costs to $i$ from her own effort. Note that these are symmetric across agents and that utility is strictly concave in own-effort. The parameter $\lambda$ determines the strength of the complementarity effect: the higher $\lambda$ is the more higher efforts by $i$'s neighbors lead to higher incentives for $i$ to increase her effort. This is the key results from Ballester et al. [2006]:

**Theorem [Ballester et al. (2006)].** *Let $\frac{\lambda}{\beta} \equiv \lambda^* < 1/\mu_1(\boldsymbol{g})$, where $\mu_1(\boldsymbol{g})$ is the largest eigenvalue of the adjacency matrix of $\boldsymbol{g}$, then we have that:*

$$\boldsymbol{e}(\boldsymbol{g}) = \frac{\alpha}{\beta} \left[ \boldsymbol{I} - \left( \frac{\lambda}{\beta} \right) \boldsymbol{g} \right]^{-1} \boldsymbol{1}$$

*where $\boldsymbol{I}$ is the identity matrix and $\boldsymbol{1}$ is a vector of 1s.*

This closed-form solution relates the effort each agent exerts at Nash to the overall network structure. Intuitively, an agent equilibrium's effort increases in how well-connected she is, how well-connected her neighbors are, how well-connected her neighbors' neighbors are, and so on... The complementarity effect determines how much weight the connectedness of neighbors, neighbors' neighbors, etc. has on an agent's equilibrium play. The Nash equilibrium is unique as long as the complementarity effect is below a threshold, because if the complementarity is too high then the equilibrium play becomes infinite effort.

Bramoullé et al. [2014] extend the result in Ballester et al. [2006] and they show that this is a potential game. Using standard results from the potential game theory literature, it follows that the unique Nash equilibrium is the equilibrium outcome of a best-reply dynamic process of agents playing this game on a network. This suggests that

[5] Ballester et al. [2006] consider weighted networks and a more general theoretical framework. Here we limit the exposition to a simplified model that suffices for our purposes and facilitates understanding.

if subjects are randomly matched to play a series of one-shot versions of the game in the lab then over time if they best-respond to the other subjects' choices they will converge to the Nash equilibrium play.

Despite the abundance of theoretical research in the economics of networks, there is still a dearth of work that shows that the intricate details of the structure of the social network provides insights on behavior above and beyond peer effects. The uniqueness of the solution and the richness of the dependence of an agent's effort on the network structure makes Ballester et al. [2006] an ideal setting to test the impact of social network structure on behavior. It also provides an ideal setting to investigate whether Nash-based results that tie network structure to outcomes are behaviorally relevant, or whether the complexity of processing network information may lead to outcomes different from the Nash predictions.

# 3 Experiment: Design and hypotheses

The experiment was conducted using UbiquityLab, a novel platform we developed to run online experiments that allows a large number of subjects in different geographical locations to play a game by interacting in real-time. The platform is composed of server-side components written in Python and client-side libraries written in JavaScript, and its modular design allows experimenters to conduct online almost every type of experiment currently run in the lab. In order to create an experiment, the researcher creates several minimal Python modules which are plugged into the servers using an API. The platform is designed to operate independently, although in the experiment in this paper we utilize *Amazon Mechanical Turk (AMT)* as a source of subjects. Appendix A contains further details on the platform.

In total, 234 subjects participated in the experiment. We ran a total of 11 meta-sessions, and the number of subjects per meta-session varied from 12 to 60 so there were up to 5 sessions in the same meta-session. Each session lasted an average of 31 minutes, and there were 6 sessions for each treatment. The participants were recruited using a 5-minute survey on *AMT*, and section 4.1 shows that they span a broad range of socio-economic characteristics. The average earnings for a subject was approximately $4.79, which is inclusive of a $1 show-up fee.[6] The experimental instructions are available in appendix C.

After clicking on a URL provided on *AMT*, subjects enter the platform and start reading the instructions. The subjects are obliged to stay a minimum amount of time on each instruction page, and they can go back to previous instruction pages at any time.

[6]These earnings are above the average per hour earnings for a worker on *AMT* and the US minimum wage. We also estimate that they are comparable to average per hour earnings in lab experiments if one takes into account that online experiments eliminate many of the activities that require subjects' time such as getting to and leaving the lab location, waiting for late participants, payment of earnings, etc.

The first part of the instructions describes the general set-up and how to pick a level of effort[7]. The second part of the instructions includes an in-depth description of how points are computed depending on their decisions and the decisions of their neighbors, as well as two illustrative examples. At the end of the instructions participants take a quiz that tests their understanding, and they are told in advance that it is necessary to answer the quiz questions correctly to proceed to the experiment. The quiz has 3 questions that test subjects' understanding of how the points are computed. There is complete anonymity throughout the experiment.

Before the beginning of the experiment each participant plays a trial on a 5-node network that consists of 3 rounds. In each round of the trial each participant is assigned to the same node of the trial network, and the effort of the neighbors are randomly chosen by the computer. The trial allows the participants to become acquainted with the interface of the experiment, and the time they have to make a decision (20 seconds). In the left part of the screen they see the network with a position highlighted in blue that denotes the node they have been assigned to, and the countdown timer. The top right part of the screen tells participants the one-letter ID of the node they have been assigned to and the IDs of the neighboring nodes in this round. Below there is a slider that allows them to pick their effort level for this round. Once everyone has selected an effort level, the results of the round are displayed in the bottom-right part of the screen.

Participants play 40 rounds of the game on the same network structure, which is always displayed throughout the experiment. In a round each subject is randomly assigned to a different node of the network[8], and picks a level of effort which can be any integer in the $[0, 100]$ range. Participants win points in each round depending on their own and their neighbors' choice of efforts, according to the formula in (1) with values of $(\alpha, \beta, \lambda)$ as below. At the end of a round each participant is reminded of the network position she was assigned to, the effort level she chose, and who were her neighbors; she also receives information about the effort choices of her neighbors and the number of points that she won.

At the end of the experiment the subjects fill in a questionnaire with standard socio-economic questions, and they take a Holt and Laury [2002] incentivized test to elicit their risk preferences. At the end of the questionnaire, we randomly pick $X$ rounds and convert the sum of points won in those rounds to the final earnings on top of the fixed \$1 fee. See appendix C for a copy of the instructions, the quiz, and the post-experimental questionnaire.

There are 4 treatments in the experiment that differ in the network that subjects are

[7]In the experiment we used the wording "level of activity" rather than "effort" to avoid priming.

[8]In networks $\mathbf{g}^{15}$ and $\mathbf{g}^{21}$ we introduce "mirror" positions so there are 9 subjects playing in $\mathbf{g}^{15}$ and 12 subjects playing in $\mathbf{g}^{21}$. In the visual interface, a subject can be assigned to any position, but when a subject is assigned to a node type $P$, $H$ or $B$ in, say, the left cluster then the subject's play counts also for the mirror node in the right cluster. This facilitates learning on node types $F$ and $C$ that are rare in the networks. Note that in $\mathbf{g}^{15}$ the $H$ nodes which are neighbors of $F$ are not mirror nodes.

exogenously assigned to. Figure 1 illustrates the 4 networks used in the experiment: a circle network $\mathbf{g}^{\bigcirc}$ with 9 nodes, a wheel network $\mathbf{g}^{\otimes}$ with 9 nodes, a 15-node network $\mathbf{g}^{15}$ with 2 completely connected clusters joined by a single node, and a 21-node network $\mathbf{g}^{21}$ with 2 completely connected clusters joined by a chain of 3 nodes. In treatments $\mathbf{g}^{\bigcirc}$, $\mathbf{g}^{\otimes}$ and $\mathbf{g}^{15}$ we chose parameter values of $\alpha = 20$, $\beta = 1$ and $\lambda = 0.1$. In treatment $\mathbf{g}^{21}$ we chose parameter values of $\alpha = 20$, $\beta = 1$ and $\lambda = 1/6$. The choice of different values of the parameters for $\mathbf{g}^{21}$ is inconsequential for our analysis because we investigate how subjects behave depending on their position within the $\mathbf{g}^{21}$ network, rather than making comparisons with subjects' behavior in the other networks.

The networks in figure 1 have labels that indicate different "types" of nodes.[9] Note that we are using a definition of "type" for expository purposes: it does not coincide with the predicted Nash equilibrium play as we want to allow for potential deviations from Nash play for network positions that are equivalent in terms of Nash play but differ in terms of structural features irrelevant to Nash play. An example is the focal node $F$ and the peripheral nodes $P$ in $\mathbf{g}^{15}$: the Nash equilibrium play is the same, but the nodes are classified as different types because, e.g., the focal position of $F$ may lead subjects to treat this node differently from the peripheral nodes.

Table 1: Equilibrium Nash and collaborative norm predictions for all treatments.

| **Network** | **Position(s)** | **Nash** | **Collaborative**† | **%Δ Welfare**‡ |
|---|---|---|---|---|
| $\mathbf{g}^{\bigcirc}$ | any | 25 | 33 | 7% |
| $\mathbf{g}^{\otimes}$ | H | 44 | 100 | 61% |
| | P | 31 | 66 | |
| $\mathbf{g}^{15}$ | H | 38 | 100 | 108% |
| | F, P | 28 | 60 | |
| $\mathbf{g}^{21}$ | H | 77 | 100 | 61% |
| | B | 78 | 100 | |
| | C | 38 | 87 | |
| | F, P | 33 | 100 | |

† Effort levels for the collaborative norm that leads to the highest welfare for the group. Note that we constrain types of nodes that have the same Nash equilibrium effort to make the same effort in the collaborative norm as well. ‡ % increase in welfare from Nash equilibrium for the group if everyone chooses effort levels according to the collaborative norm.

The first and second columns of Table 3 summarize the different networks and positions respectively. The third column lists the Nash equilibrium effort levels rounded at the nearest integer. For instance, the equilibrium effort level $e^*_H(\mathbf{g}^{\otimes})$ for the hub node $H$

[9] These are of course different from the labels that were displayed to subjects: in the experiment each node had a one-letter label inside the node and these labels were fixed for all sessions in a treatment.

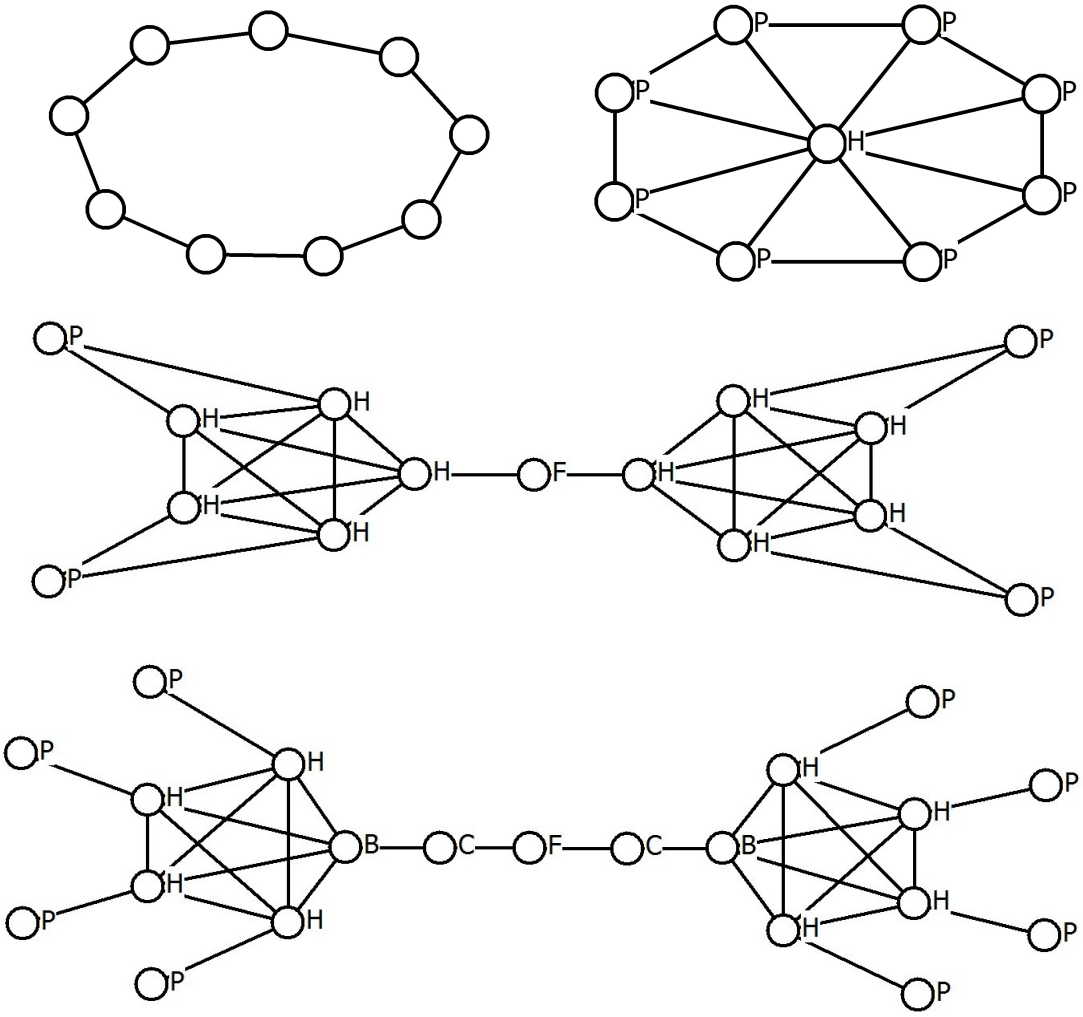



Figure 1: Networks used in the experiment. Top left: circle network $\mathbf{g}^{\bigcirc}$ with 9 nodes. Top right: wheel network $\mathbf{g}^{\otimes}$ with 9 nodes. Middle: 15-node network $\mathbf{g}^{15}$ with 2 completely connected clusters joined by a single node. Bottom: 21-node network $\mathbf{g}^{21}$ with 2 completely connected clusters joined by a chain of 3 nodes.

in the wheel network $\mathbf{g}^{\otimes}$ is 44 and the equilibrium effort level $e_P^*(\mathbf{g}^{\otimes})$ for the peripheral node $P$ is 31. However, Nash equilibrium play does not maximize the welfare of the group, defined as the sum of the utilities of all the subjects. In all 4 networks it is possible for subjects to increase their earnings by coordinating on a *collaborative norm*: choosing an effort higher than the Nash equilibrium effort to get a higher payoff as long as everyone else follows the norm as well. The fourth column lists the effort levels for the

most efficient collaborative norm in terms of maximizing the group's welfare.[10] The last column in Table 3 shows that there are very large welfare gains for the group if everyone follows the collaborative norm. For instance, in $\mathbf{g}^{\otimes}$ the collaborative norm that achieves the highest group welfare leads to a 61% increase in group welfare compared to the case where everyone plays Nash.

The first question that we want to investigate is whether subjects are able to coordinate on a collaborative norm. In any collaborative norm the subjects exert higher effort than equilibrium play in each network position so a basic test of the presence of a collaborative norm is that the observed effort levels the subjects converge to in the experiment are higher than the Nash predictions. This hypothesis leads to 11 inequalities that we can test for in the four networks.

*H1: Subjects converge to a collaborative norm with effort levels higher than Nash play. The specific hypotheses to test are:*

*(i)* $e_F(\boldsymbol{g}^{\bigcirc}) > 25$

*(ii)* $e_H(\boldsymbol{g}^{\otimes}) > 44$ *and* $e_P(\boldsymbol{g}^{\otimes}) > 31$

*(iii)* $e_H(\boldsymbol{g}^{15}) > 38$, $e_P(\boldsymbol{g}^{15}) > 28$ *and* $e_F(\boldsymbol{g}^{15}) > 28$

*(iv)* $e_B(\boldsymbol{g}^{21}) > 78$, $e_H(\boldsymbol{g}^{21}) > 77$, $e_C(\boldsymbol{g}^{21}) > 38$, $e_P(\boldsymbol{g}^{21}) > 33$ *and* $e_F(\boldsymbol{g}^{21}) > 33$

In case subjects are unable to coordinate on a collaborative norm, a potential fall back option is to play the Nash equilibrium level of effort for each network position. This generates 11 point predictions that we can test for in the 4 different network structures.

*H2: Subjects converge to Nash play. The specific hypotheses to test are:*

*(i)* $e_F(\boldsymbol{g}^{\bigcirc}) = 25$

*(ii)* $e_H(\boldsymbol{g}^{\otimes}) = 44$ *and* $e_P(\boldsymbol{g}^{\otimes}) = 31$

*(iii)* $e_H(\boldsymbol{g}^{15}) = 38$ *and* $e_P(\boldsymbol{g}^{15}) = e_F(\boldsymbol{g}^{15}) = 28$

*(iv)* $e_B(\boldsymbol{g}^{21}) = 78$, $e_H(\boldsymbol{g}^{21}) = 77$, $e_C(\boldsymbol{g}^{21}) = 38$ *and* $e_P(\boldsymbol{g}^{21}) = e_F(\boldsymbol{g}^{21}) = 33$

However, the complexity of the network structure may make it very challenging for subjects to converge to Nash equilibrium play, either through learning or by figuring out what the equilibrium is. An intuitive way to reduce this complexity is to focus solely on the local, rather than the global, structure of the network. The degree, or number

[10] Note that coordination on the collaborative norm that achieves the highest welfare may require subjects to take losses when assigned to certain positions in the network, which would naturally make the norm more difficult to sustain. However, in all 4 networks there are collaborative norms such that subjects achieve higher payoffs than Nash in every network position.

of connections, of a node is a very prominent and easily distinguishable local feature of the structural position of a node in the network. Despite being a rather simple metric, it is also highly correlated with Bonacich centrality, which determines Nash play. The fact that degree is a very intuitive metric and that it is highly correlated with Bonacich makes it a natural way to investigate the role of the local structure of the network: subjects who choose their efforts depending on the local structure would pick higher efforts the higher is the degree of the node they are assigned to.

A drawback of the high correlation between degree and Bonacich centrality is that in most networks subjects whose play is determined by the local structure of the network will be undistinguishable from subjects playing Nash. However, this is not the case for $\mathbf{g}^{21}$ with the value of $\lambda$ that we have chosen. As Table 3 shows, the equilibrium Nash prediction is $e^*_C(\mathbf{g}^{21}) > e^*_F(\mathbf{g}^{21}) = e^*_P(\mathbf{g}^{21})$. From figure 1 we can see that nodes $F$ and $C$ have degree 2 while node $P$ has only one connection so subjects whose play is determined by the local network structure would pick efforts $e_F(\mathbf{g}^{21}) = e_C(\mathbf{g}^{21}) > e_P(\mathbf{g}^{21})$, which is a different ranking from Nash play. This leads to a complete ranking of effort levels that we can test for in networks $\mathbf{g}^{\otimes}$, $\mathbf{g}^{15}$ and $\mathbf{g}^{21}$, and that, crucially, departs from the ranking for Nash play for 3 types of nodes in $\mathbf{g}^{21}$:

*H3: Subjects choose effort levels based on their local position in the network as captured by degree, rather than taking into account how their position relates to the overall network structure as captured by Bonacich centrality.The specific hypotheses to test are:*

*(i)* $e_H(\boldsymbol{g}^{\otimes}) > e_P(\boldsymbol{g}^{\otimes})$

*(ii)* $e_H(\boldsymbol{g}^{15}) > e_F(\boldsymbol{g}^{15}) = e_P(\boldsymbol{g}^{15})$

*(iii)* $e_H(\boldsymbol{g}^{21}) = e_B(\boldsymbol{g}^{21}) > e_F(\boldsymbol{g}^{21}) = e_C(\boldsymbol{g}^{21}) > e_P(\boldsymbol{g}^{21})$

# 4 Experiment: Results

Section 4.1 describes basic data on the sample of subjects, and section 4.2 presents the results.

## 4.1 Sample description

The experimental data contains the decisions of 234 subjects. Each subject participated in one session of one treatment so she played the game in one of the four network structures. We ran each treatment 6 times and there are 40 rounds in each game so we have a total of $9,360$ decisions or observations. The experimental data was matched with the data from the questionnaire that subjects had to fill in at the end of the experiment.

Table 2 summarizes some of the main socio-demographic characteristics of the participants and their pairwise correlations. All the participants are resident in the US: 39% are

Table 2: Summary statistics and pairwise correlations for some of the main variables.[1]

| | n | Mean | s.d. | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|---|---|
| 1. Gender[1] | 233 | 0.39 | 0.49 | 1.00 | | | | |
| 2. Age | 233 | 30.72 | 10.48 | 0.00 | 1.00 | | | |
| 3. Trust[2] | 233 | 0.52 | 0.50 | -0.04 | -0.05 | 1.00 | | |
| 4. Education [3] | 233 | 3.35 | 1.27 | -0.00 | 0.00 | 0.07 | 1.00 | |
| 5. HL[4] | 212 | 7.30 | 1.59 | −0.02 | -0.10 | 0.14* | 0.06 | 1.00 |

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. 1. Female= 1. 2. General trust question with 0 =“Need to be very careful” and 1 =“Most people can be trusted”; 3. See the caption of Figure 2 for the coding; 4. Holt and Laury [2002]'s risk attitude test: 20 participants are excluded because they made at least one inconsistent choice (i.e. multiple switching points) and 1 participant is excluded because he/she had no switching point. Note that one participant did not complete the questionnaire.

female and the average age is 30.7 years old. We asked subjects the standard interpersonal trust question from the World Values Survey ($WVS$) and found that 52% believe that others can be trusted, which is higher than the average value from the $WVS$ survey of the US population. The level of trust is not correlated with gender or age.

Figure 2 illustrates in more detail one of the advantages of online experiments vis-à-vis lab studies: a heterogeneous subject pool that is more representative of the general population than the student populations that are typical of most laboratory studies. Figure 2(a) shows that subjects belong to different age groups ranging from a minimum of 18 to a maximum of 70 years old. Moreover, there is a significant heterogeneity in terms of the education level of the participants. Figure 2(b) shows that 76.8% of participants have some kind of college degree as the highest level of education attained, 6.9% only have a high school diploma, and 16.3% have a master, professional degree or PhD.

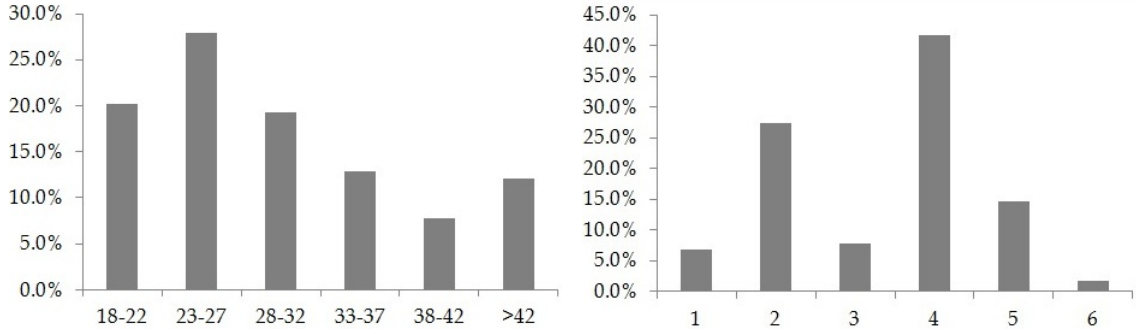



Figure 2: Percentages of subjects of different age groups (left) and education levels (right) for 233 out of 234 participants. Classification of education levels: 1 = high school, 2 = some college, 3 = 2-year college, 4 = 4-year college, 5 = master, 6 = professional degree (e.g. MD) or PhD.

After the questionnaire, subjects took the classic Holt and Laury [2002]'s risk attitude test. For each of 10 scenarios they had to pick between a safe and a risky lottery. In the early scenarios the safe lottery gives a higher expected payment, while in the late scenarios it is the other way around. A risk-neutral participant would switch from the safe to the risky lottery at scenario 5. The participants are on average risk-averse because the mean switching point is 7.3. There are 4 (1.9%) participants who switch at scenario 4 and therefore are slightly risk-seeking, and 29 (13.7%) participants who are risk-neutral. There is a significant correlation between risk aversion and trust: participants who display a high level of trust tend to be more risk-averse. There is no significant correlation of risk aversion with gender, age or education.

## 4.2 Results

In this section we test the hypotheses formulated in section 3 using the experimental data. Table 3 shows the average and median effort in the last 10 rounds for each type of node, and figures 3 and 4 show the distribution of efforts for each type of node in the last 10 rounds.

Our first finding is that subjects are able to coordinate on a collaboration norm in the simple $\mathbf{g}^{\bigcirc}$ and $\mathbf{g}^{\otimes}$ networks. In the $\mathbf{g}^{\bigcirc}$ network, subjects' average effort level in the last 10 rounds is equal to 29.2, which is significantly higher (Wilcoxon signed rank test, hereafter WSR, $p = 0.02$) than the Nash prediction $e^*(\mathbf{g}^{\bigcirc}) = 25$. As we saw in Table 3, the collaborative norm that would lead to the highest welfare for the group requires an effort level equal to 33 so subjects converge to an average effort level that is in between Nash and the most efficient collaborative norm. The top left panel of figure 3 shows the distribution of effort levels in the last 10 rounds: the large majority of subjects choose efforts in between the Nash prediction and the most efficient collaborative norm. There is some loose analogy between this type of collaborative norm and the contribution norms observed in public good experiments with quadratic payoffs: Laury and Holt [2008] find levels of collaboration that are similar to what we find in the $\mathbf{g}^{\bigcirc}$ treatment.[11]

The collaborative norm is still present in the $\mathbf{g}^{\otimes}$ network, but it is weaker. The average effort level in the last 10 rounds in the $H$ type node is 51, which is qualitatively higher than the Nash prediction $e^*_H(\mathbf{g}^{\otimes}) = 44$. Similarly, the average effort level in the last 10 rounds in the $P$ type node is 34.9, which is significantly higher (WSR, $p = 0.03$) compared to the Nash prediction $e^*_P(\mathbf{g}^{\otimes}) = 33$. However, the collaborative norm that we observe is extracting only a small fraction of the potential gains from coordination compared to the most efficient collaborative norm which, as table 3 shows, would entail subjects in the $H$ type node choosing effort level equal to 100 and subjects in the $P$ or $F$ type nodes choosing 60. Thus, the average effort levels subjects converge to are much

[11] See, in particular, their low Nash treatment where the Nash contribution is 20% and the Pareto optimum is 80%: subjects' contributions start close to 60% and decline to just below 40%.

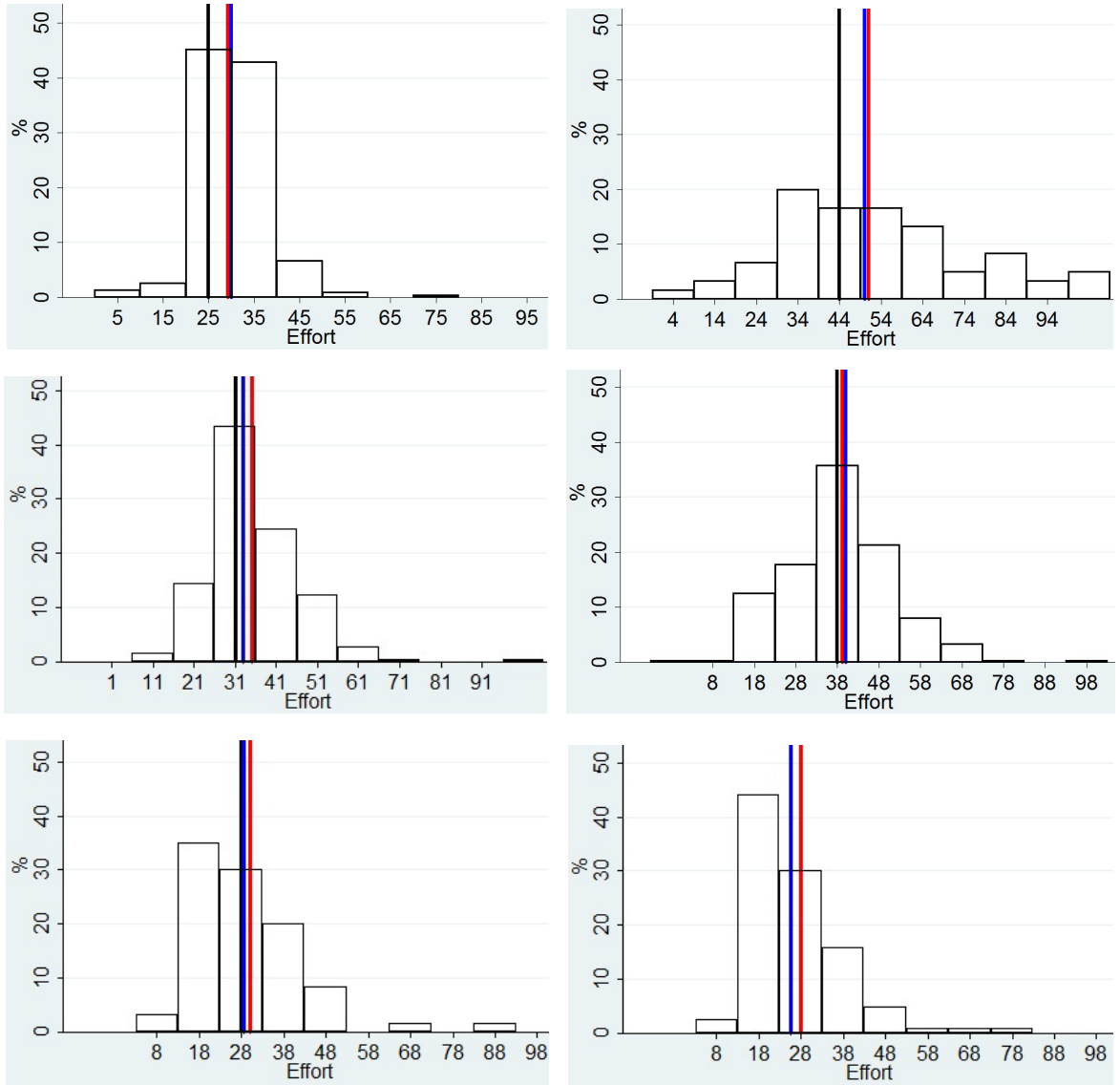



Figure 3: Distribution of efforts chosen in the last 10 rounds by type of position in networks $\mathbf{g}^{\bigcirc}$, $\mathbf{g}^{\otimes}$ and $\mathbf{g}^{15}$. Top: $\mathbf{g}^{\bigcirc}$ (left) and $H$ type in $\mathbf{g}^{\otimes}$ (right). Middle: $P$ type in $\mathbf{g}^{\otimes}$ (left) and $H$ type in $\mathbf{g}^{15}$. Bottom: $F$ (left) and $P$ (right) type in $\mathbf{g}^{15}$. Vertical lines indicate predicted Nash play (black), mean (red) and median (blue) efforts.

Table 3: Effort levels in different network positions.

| Network | Position(s) | Nash | Effort last 10 rounds | | |
|---|---|---|---|---|---|
| | | | Mean | s.d. | Median |
| $\mathbf{g}^{\bigcirc}$ | any | 25 | 29.32** | 7.63 | 30 |
| $\mathbf{g}^{\otimes}$ | H | 44 | 51.02 | 22.88 | 50 |
| | P | 31 | 34.89** | 10.98 | 33 |
| $\mathbf{g}^{15}$ | H | 38 | 39.15 | 12.63 | 40 |
| | F | 28 | 29.92 | 13.62 | 28.5 |
| | P | 28 | 27.88 | 11.45 | 25.5 |
| $\mathbf{g}^{21}$ | H | 77 | 78.95 | 24.01 | 88 |
| | B | 78 | 79.35 | 24.28 | 86 |
| | C | 38 | 40.79 | 16.00 | 40 |
| | F | 33 | 38.95 | 19.86 | 34 |
| | P | 33 | 34.63 | 18.25 | 30 |

Summary statistics for effort levels in the different network positions for the last 10 rounds of play. Significance levels refer to the Wilcoxon signed-rank test on aggregated data at the session level for the last 10 rounds ($^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$).

closer to Nash than to the most efficient collaborative norm. An interpretation is that the $\mathbf{g}^{\otimes}$ network introduces asymmetry across nodes, which makes the coordination on a collaborative norm more challenging. Continuing with the loose analogy with public good game experiments, this is in agreement with the finding that contribution levels decrease in public good games where there are asymmetries across participants (see, e.g., Rapoport and Suleiman [1993]).

*Result 1: In the $\boldsymbol{g}^{\bigcirc}$ network, subjects establish and maintain a collaborative norm that entails choosing efforts higher than the Nash prediction. The collaborative norm is still present, but it is weaker, in the asymmetric $\boldsymbol{g}^{\otimes}$ network. The norm disappears in the more complex and asymmetric $\boldsymbol{g}^{15}$ and $\boldsymbol{g}^{21}$ networks.*

If subjects are unable to sustain a collaboration norm in the more complex and asymmetric networks, a potential fall back option is to play the Nash equilibrium effort levels. At the same time, the set-up of this game is rather complex given the rich structure of the $\mathbf{g}^{15}$ and $\mathbf{g}^{21}$ networks as well as the number of discrete effort levels subjects can choose from, so it is not an easy task to figure out or learn the equilibrium.

Our second finding is that subjects converge to the Nash equilibrium predictions in the more complex $\mathbf{g}^{15}$ and $\mathbf{g}^{21}$ networks. Table 3 shows that in $\mathbf{g}^{15}$ the average effort in the last 10 rounds is statistically indistinguishable from the Nash predictions independently of the type of position. The average effort by subjects assigned to the $H$ node is 39.2,

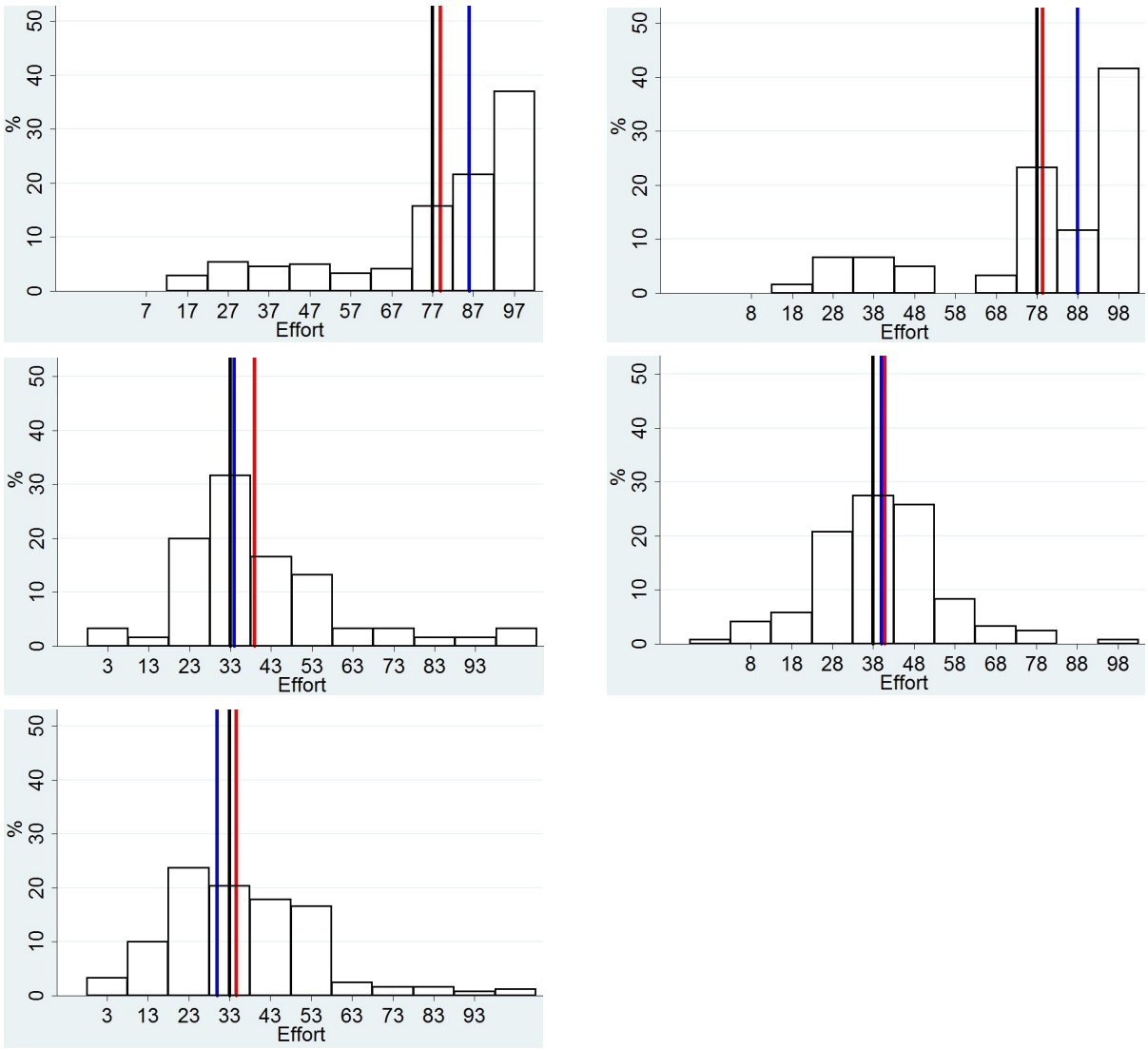



Figure 4: Distribution of efforts chosen in the last 10 rounds by type of position in network $\mathbf{g}^{21}$. Top: $H$ (left) and $B$ (right) type. Center: $F$ (left) and $C$ (right) type. Bottom: $P$ type. Vertical lines indicate predicted Nash play (black), mean (red) and median (blue) efforts.

and we cannot reject the hypothesis that this is the same as the Nash effort $e^*_H(\mathbf{g}^{15}) = 38$. Similarly, the average efforts in the last 10 rounds in the $F$ and $P$ positions are 29.9 and 27.9 respectively, and we cannot reject hypothesis $H2(iii)$ that they are the same as the Nash predictions $e^*_F(\mathbf{g}^{15}) = e^*_P(\mathbf{g}^{15}) = 28$. The center right and bottom panels of figure 3 show the distribution of efforts in the last 10 rounds in the $H$, $F$ and $P$ type of nodes respectively: efforts are approximately normally distributed with mean/median at the Nash equilibrium prediction.

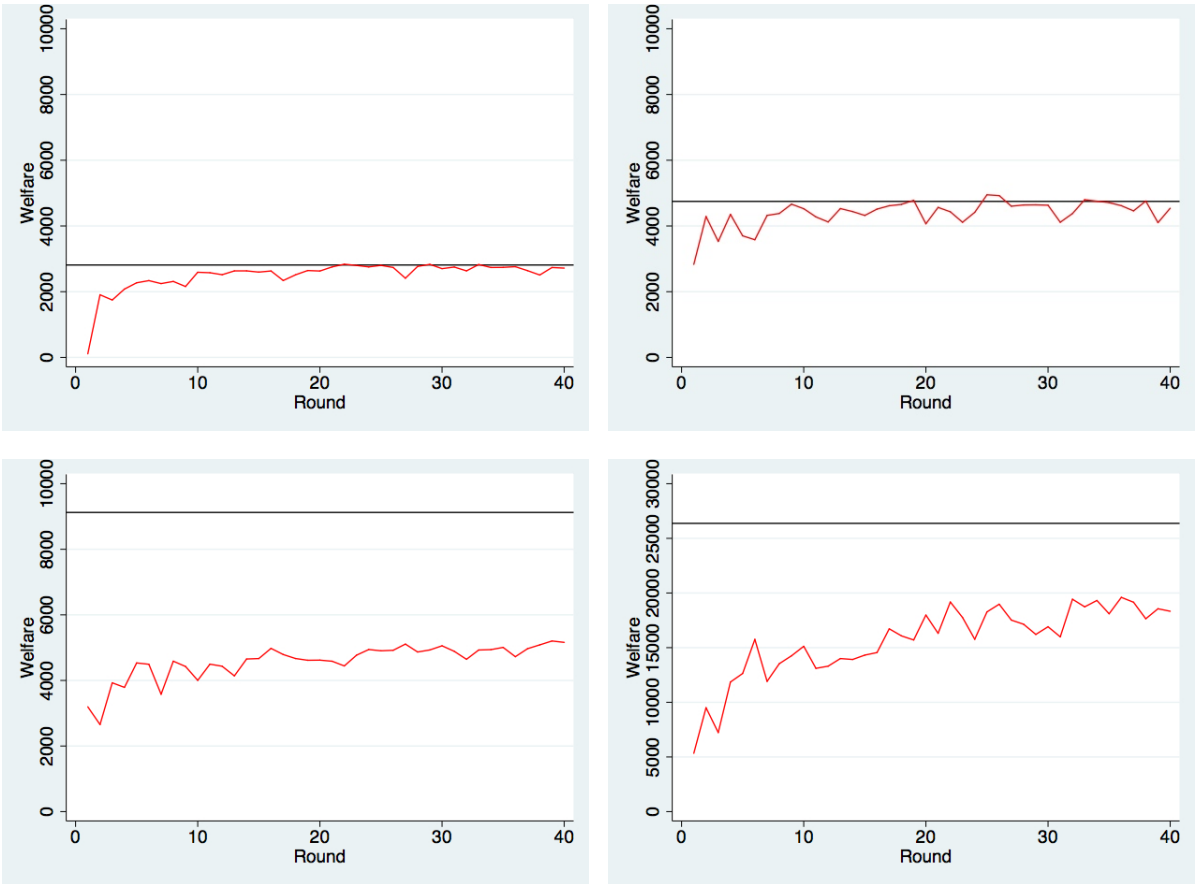



Figure 5: Evolution of welfare level during the 40 rounds. Top: $\mathbf{g}^{\bigcirc}$ (left) and $\mathbf{g}^{\otimes}$ (right) treatment. Bottom: $\mathbf{g}^{15}$ (left) and $\mathbf{g}^{21}$ (right) treatment. Lines indicate predicted Nash (black) and mean (red) welfare levels.

The Nash point predictions in $H2(iv)$ for the $\mathbf{g}^{21}$ network are also validated by the data. Table 3 shows that the average efforts in the last 10 rounds in the $B$ and $H$ positions are 79.4 and 79 respectively, and we cannot reject hypothesis $H2(iv)$ that they are the same as the Nash predictions $e^*_B(\mathbf{g}^{21}) = 78$ and $e^*_H(\mathbf{g}^{21}) = 77$. Similarly, the average efforts in the last 10 rounds in the $C$ and $P$ positions are 40.8 and 34.6 respectively, and we cannot

reject hypothesis $H2(iv)$ that they are the same as the Nash predictions $e^*_C(\mathbf{g}^{21}) = 38$ and $e^*_P(\mathbf{g}^{21}) = 33$. The Nash prediction for type $F$ node is also validated by the data, although the average effort by subjects in the last 10 rounds is 39 which is qualitatively higher than the Nash prediction $e^*_F(\mathbf{g}^{21}) = 33$. Figure 4 shows the distribution of efforts in the last 10 rounds. They tend to be approximately normally distributed with mean/median coinciding with the Nash equilibrium prediction, aside for position $F$ where the mean is qualitatively higher than the Nash prediction. In the $H$ and $B$ type nodes, there is a fraction of subjects that pick efforts close to or equal to 100, possibly due to the vicinity of the Nash prediction to the upper boundary of the strategy space.

A consequence of the subjects' inability to establish a collaborative norm in the $\mathbf{g}^{15}$ and $\mathbf{g}^{21}$ networks is that the aggregate welfare is, in comparative terms, lower than in the $\mathbf{g}^{\bigcirc}$ and $\mathbf{g}^{\otimes}$ networks relative to the theoretical aggregate welfare achievable if everyone were to play Nash. Figure 5 shows the evolution of the aggregate welfare, i.e. the sum of subjects' payoffs, for each of the 4 networks. In the $\mathbf{g}^{\bigcirc}$ and $\mathbf{g}^{\otimes}$ networks, subjects converge to an aggregate welfare that is indistinguishable from the theoretical aggregate welfare if everyone played Nash. On the other hand, the aggregate welfare in the $\mathbf{g}^{15}$ and $\mathbf{g}^{21}$ networks is significantly lower than the theoretical aggregate welfare if everyone played Nash. Notice that the presence of variability in subjects' play lowers aggregate welfare so the average aggregate welfare is not above the one achievable at Nash in the $\mathbf{g}^{\bigcirc}$ and $\mathbf{g}^{\otimes}$ networks despite the establishment of a collaboration norm, and, similarly, it is not equal to the one achievable at Nash in the $\mathbf{g}^{15}$ and $\mathbf{g}^{21}$ networks despite the subjects converging to Nash play on average.

*Result 2: In the $\boldsymbol{g}^{15}$ and $\boldsymbol{g}^{21}$ networks, subjects converge to the Nash point predictions on all node types.*

Our third finding is that there is evidence that subjects focus on the local, rather than the global, features of the network to make their decisions. The first piece of supporting evidence is that the ranking of subjects' effort decisions in different types of nodes matches *all* the predictions in $H3$. Table 4 shows the average difference in effort levels of subjects across the different types of nodes in the $\mathbf{g}^{\otimes}$, $\mathbf{g}^{15}$ and $\mathbf{g}^{21}$ networks. As we have hypothesized in section 3, subjects who choose their efforts depending on the local structure would pick higher efforts the higher is the degree of the node they are assigned to. In the $\mathbf{g}^{\otimes}$ and $\mathbf{g}^{15}$ networks, the predictions in $H3(i)$ and $H3(ii)$ of subjects playing based on the local structure match the Nash predictions, and Table 4 shows that the data validates these predictions. Similarly, in the $\mathbf{g}^{21}$ network, subjects in the $B$ and $H$ node types choose effort levels higher than subjects in the other node types in agreement with both the Nash predictions and $H3$. However, subjects who play according to the degree will deviate from Nash play for moderately high levels of the complementarity effect in specific networks, and we designed the $\mathbf{g}^{21}$ treatment to examine such a case.

In the $\mathbf{g}^{21}$ network, there are three node types where the Nash predictions ranking of effort levels differs from the ranking in $H3$, and the data fully validates the predictions

Table 4: Differences in average effort across different positions for all the data.

| | | $\mathbf{g}^{\otimes}$ | $\mathbf{g}^{15}$ | | $\mathbf{g}^{21}$ | | | |
|---|---|---|---|---|---|---|---|---|
| | | P | F | P | H | P | C | F |
| $\mathbf{g}^{\otimes}$ | H | $14.0^{***}$ | | | | | | |
| $\mathbf{g}^{15}$ | H | | $9.2^{**}$ | $9.7^{**}$ | | | | |
| | F | | | 0.5 | | | | |
| $\mathbf{g}^{21}$ | B | | | | 0.6 | $37^{***}$ | $29.8^{***}$ | $32.5^{***}$ |
| | H | | | | | $37.6^{***}$ | $30.4^{***}$ | $33.2^{***}$ |
| | P | | | | | | $-7.2^{**}$ | $-4.5^{*}$ |
| | C | | | | | | | 2.8 |

The analysis is based on aggregated data at the session level, and each entry is the difference between the average effort in the row position and the average effort in the column position. Significance levels refer to the Mann-Whitney test. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

generated by play based on local network structure. Subjects playing according to the degree would pick the same effort level in node types $F$ and $C$, which have degree 2, and they would pick a higher effort level in node $F$ compared to the $P$ node type that has degree 1, while the Nash predictions are $e^*_C(\mathbf{g}^{21}) > e^*_F(\mathbf{g}^{21})$ and $e^*_F(\mathbf{g}^{21}) = e^*_P(\mathbf{g}^{21})$. Table 4 validates the predictions in $H3(iii)$: the effort levels of subjects in the $F$ and $C$ node types are indistinguishable, and subjects choose a higher effort level in node $F$ compared to the $P$ node type (Mann-Whitney, $p = 0.068$). Moreover, we also cannot reject the hypothesis that subjects' play in node type $F$ is the same as the Nash prediction $e^*_C(\mathbf{g}^{21}) = 38$ for $C$.

The second piece of supporting evidence is a panel data analysis of the determinants of subjects' efforts in Table 5. Specifications (1)-(2) analyze the full data for the $\mathbf{g}^{21}$ treatment: the common regressors to all the specifications are subjects' gender, age, trust and risk aversion as captured by the risk elicitation test in the post-experimental questionnaire. In specification (1) we add the degree of the node the subject was assigned to as an additional regressor, and we find that it is highly significant. In (2) we repeat the same exercise using Bonacich centrality instead of degree, and, similarly, we find that it is highly significant. Specification (2) includes both degree and Bonacich centrality as regressors and both remain highly significant, which makes it challenging to distinguish between the two metrics.

Bonacich centrality may still be a predictor of subjects' decisions even though the subjects choose efforts according to their degree, because it is highly correlated to degree for the majority of node positions. In the case of $\mathbf{g}^{21}$, node types $H$ and $B$ have both high degree and centrality. In specifications (3)-(5) we repeat the panel data analysis excluding from the sample node types $H$ and $B$, and keeping node types $F$, $C$ and $P$

Table 5: Determinants of subjects' efforts.

| | $\mathbf{g}^{\otimes}$ | $\mathbf{g}^{15}$ (all) | $\mathbf{g}^{15}$ ($\{F, P\}$) | $\mathbf{g}^{21}$ (all) | $\mathbf{g}^{21}$ ($\{F, C, P\}$) | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Degree | 2.828*** | 2.973*** | | 4.353** | 4.946*** | | |
| | (0.176) | (0.212) | | (1.411) | (1.355) | | |
| Bonacich | | | | 10.446*** | 6.643 | | |
| | | | | (2.413) | (5.294) | | |
| $F$ position | | | | | | -1.860 | |
| | | | | | | (1.482) | |
| $P$ position | | | -0.586 | | | -6.806*** | -4.946*** |
| | | | (1.118) | | | (1.047) | (1.355) |
| $C$ position | | | | | | | 1.860 |
| | | | | | | | (1.482) |
| Round | -0.052* | -0.063 | -0.201*** | 0.199*** | -0.132*** | -0.132*** | -0.132*** |
| | (0.024) | (0.026) | (0.457) | (0.034) | (0.039) | (0.039) | (0.039) |
| Age | 0.117 | -0.046 | 0.034 | -0.357* | -0.212 | -0.212 | -0.212 |
| | (0.087) | (0.074) | (0.092) | (0.144) | (0.129) | (0.129) | (0.129) |
| Gender | 2.011 | 2.809 | 4.377* | 3.309 | 4.895 | 4.895 | 4.895 |
| | (1.702) | (1.757) | (2.183) | (2.853) | (2.549) | (2.549) | (2.549) |
| Trust | 1.935 | -0.320 | 1.492 | 4.546 | 3.545 | 3.545 | 3.545 |
| | (1.702) | (1.721) | (2.145) | (2.971) | (2.652) | (2.652) | (2.652) |
| Risk | 0.214 | 1.667 | 0.495 | 7.123* | 1.702 | 1.702 | 1.702 |
| | (1.745) | (2.158) | (2.682) | (3.331) | (2.979) | (2.979) | (2.979) |
| Sessions | Y | Y | Y | Y | Y | Y | Y |
| Constant | 16.771*** | 27.878*** | 27.878*** | 14.143 | 24.443* | 47.088*** | 45.228*** |
| | (5.053) | (4.961) | (6.125) | (7.694) | (10.185) | (6.630) | (6.679) |
| Obs. | 2160 | 2160 | 720 | 2840 | 1654 | 1654 | 1654 |
| Subjects | 54 | 54 | 54 | 72 | 72 | 72 | 72 |
| Wald $\chi^2$ | 292.37 | 242.96 | 41.44 | 2210.78 | 75.78 | 75.78 | 75.78 |
| Prob> $\chi^2$ | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

Tobit panel estimation with random effects at subject level and session fixed effects.
Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

where the degree and Bonacich centrality metrics differ. Specification (3) has degree as a regressor, and it is highly significant. Similarly, (4) has Bonacich centrality instead of degree, and it is also highly significant. However, in (5) we include both degree and Bonacich as regressors and we find that degree remains highly significant while Bonacich is completely insignificant. This is evidence that subjects are basing their decision on their position in the local rather than the global network structure: in a network where degree and Bonacich are not perfectly correlated, degree is highly predictive of subjects' choices, while Bonacich centrality has no predictive power once the degree is included as a regressor.

There are two separate effects which lead degree to be a better predictor of subjects' choices than Bonacich centrality. Specification (6) repeats the panel data analysis with dummies for the $F$ and $P$ positions, and with the $C$ position as a baseline. The $F$ position dummy is not significant, which indicates that the first reason degree is more predictive is that there is no difference in effort levels between the $F$ and $C$ positions as predicted by degree and in contrast to the $e_C(\mathbf{g}^{21}) > e_F(\mathbf{g}^{21})$ Bonacich prediction. Specification (7) repeats the panel data analysis with dummies for the $C$ and $P$ positions. The $P$ position dummy is negative and highly significant, which indicates that the second reason degree is more predictive is that there is a difference in effort levels between the $F$ and $P$ positions as predicted by degree and in contrast to the $e_P(\mathbf{g}^{21}) = e_F(\mathbf{g}^{21})$ Bonacich prediction.

A third piece of supporting evidence for the local features of the network structure as the main determinant of subjects' choices is qualitative in nature. In the post-experimental questionnaire, we asked subjects the following open-ended question: "Please briefly describe how you picked your activity level during the experiment." Subjects were required to input an answer in a text box, and we employed an RA to categorize the answers. Out of the subjects who answered the question[12] for the $\mathbf{g}^{15}$ and $\mathbf{g}^{21}$ networks (89.7%, 113 out of 126), two thirds of them (64.6%) replied that they chose the effort depending on the number of connections of the node they were assigned to. This was the most popular answer, followed by "trial and error" (14.2%), "maximize own profits" (9.7%), and "depending on what others were choosing" (7.1%). No subject mentioned the number of connections of their neighbors in the answer, which is an important component of Bonacich centrality.

*Result 3: We find qualitative and quantitative evidence that subjects base their choices on the local rather than the global structure of the network. In all networks, subjects' efforts are ranked according to the degree of the node they are assigned to. In the node types of the $\boldsymbol{g}^{21}$ network where the degree has different predictions than Nash, the degree is highly predictive of subjects' behavior while Bonacich centrality is insignificant. Finally, the degree of the node they are assigned to is subjects' most popular response to an open-ended question of how they make their decisions.*

[12]It was compulsory to answer the question, but some subjects decided to decline to provide an answer by, e.g., inputting a hyphen in the text box.

# 5 Robustness checks

This section presents two checks of the robustness of our results. Section 5.1 shows that the results are not driven by subjects having different exposure to node types due to the different number of nodes of a certain type in each network. Section 5.2 shows that the results are not affected by the specific network visualizations used in the interface of the experiment.

## 5.1 The role of experience

In each round of the experiment subjects are randomly assigned to nodes in the network. In the asymmetric $\mathbf{g}^{\otimes}$, $\mathbf{g}^{15}$ and $\mathbf{g}^{21}$ networks, there are different number of nodes of each type, which implies that subjects will have different degrees of experience on making decisions on different types of nodes. For instance, in the $\mathbf{g}^{\otimes}$ network there is one node of type $H$ and 8 nodes of type $P$ so a subject will on average be assigned to the $H$ node 4.4 times over 40 rounds and he will be assigned 35.6 times to one of the $P$ nodes. The purpose of this section is to check that these differences in experience are not major determinants of subjects' decisions.

Table 6: Differences in average effort across different positions for the third time a subject has been assigned to a given node type.

| | | $\mathbf{g}^{\otimes}$ | $\mathbf{g}^{15}$ | | $\mathbf{g}^{21}$ | | | |
|---|---|---|---|---|---|---|---|---|
| | | P | F | P | H | P | C | F |
| $\mathbf{g}^{\otimes}$ | H | 3.4 | | | | | | |
| $\mathbf{g}^{15}$ | H | | 11.8** | 9.1** | | | | |
| | F | | | −2.6 | | | | |
| $\mathbf{g}^{21}$ | B | | | | 8.4 | 42.4*** | 35.2*** | 36.8*** |
| | H | | | | | 34.0*** | 26.7*** | 28.3*** |
| | P | | | | | | −7.2 | −5.6 |
| | C | | | | | | | 1.6 |

The analysis is based on aggregated data at the session level, and each entry is the difference between the average effort in the row position and the average effort in the column position. Significance levels refer to the Mann-Whitney test. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 6 replicates the analysis in table 4, but it focuses on the subject's choice of effort the third time a subject has been assigned to a certain node type rather than aggregating the decisions over all rounds. For instance, if in the $\mathbf{g}^{21}$ network a subject has been

assigned to node type $P$ in rounds 1, 2 and 3 and it has been assigned to node type $F$ in rounds 9, 14 and 36 then the entry in the $P$-$F$ cell would be the difference between that subject's play in round 3 and 36. A comparison of the two tables shows that the results are largely unchanged. The only differences from table 4 are the lack of statistical significance in the difference between node types $H$ and $P$ in $\mathbf{g}^{\otimes}$, $P$ and $C$ in $\mathbf{g}^{21}$, and $P$ and $F$ in $\mathbf{g}^{21}$. However, the differences are qualitatively in the same direction, and in the case of node types $H$, $C$ and $P$ in $\mathbf{g}^{21}$ they also have the same magnitude so the lack of statistical significance is likely due to the higher variability of play in the earlier rounds for node type $P$, which is more frequent in $\mathbf{g}^{21}$ leading subjects to experience it for the third time in the early rounds.

Table 7: Differences in average effort between low and high round experience subjects, where low is $<= x$, high is $>= y$ and the last round played by each subject is considered. (x,y) are specified next to the node type.

| | $\mathbf{g}^{\otimes}$ | $\mathbf{g}^{15}$ | $\mathbf{g}^{21}$ | |
|---|---|---|---|---|
| | H(3,6) | F(3,6) | C(4,7) | F(2,5) |
| | High | | | |
| Low | 11.4 | 8.5 | 1.9 | 6.2 |

The analysis is based on aggregated data at the session level, and each entry is the difference between the average effort in the row position and the average effort in the column position. Significance levels refer to the Mann-Whitney test. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 6 suggests the obvious point that effort choices on node types that are rare in a network may be more sensitive to different degrees of experience than effort choices on node types that are frequent. We can probe further into this by exploiting the random differences between subjects within sessions with respect to the number of times they have been assigned to a given node type. Specifically, consider, say, node $F$ in $\mathbf{g}^{21}$ and let us divide subjects into two groups: low experience subjects have been assigned at most 2 times to node $F$ throughout the 40 rounds, while high experience subjects have been assigned at least 5 times. We can aggregate the last effort choice for low and high experience subjects respectively and test whether there is a significant difference between the two: a difference would indicate that the randomly determined degree of experience has an effect on the results. Table 7 conducts this exercise for node types $H$ in $\mathbf{g}^{\otimes}$, $F$ in $\mathbf{g}^{15}$, and $F$ and $C$ in $\mathbf{g}^{21}$. There is no statistically significant difference between low and high experience subjects for any of the node types suggesting that the 40 rounds duration of the experiment is sufficient to ensure that even low experience subjects have enough time to learn how to choose efforts in the node types they are only assigned to a few times.

Table 8: Differences in average effort within high experience subjects. Low ($= x$) and high ($= y$) are equal to the $x$th and last time a subject has been assigned to a given node type respectively. (x,y) are specified next to the node type.

| | $\mathbf{g}^{\otimes}$ | $\mathbf{g}^{15}$ | $\mathbf{g}^{21}$ | |
|---|---|---|---|---|
| | H(3,6) | F(3,6) | C(4,7) | F(2,5) |
| | High | | | |
| Low | $-0.5$ | $-1.9$ | $-1.4$ | 8.0 |

The analysis is based on aggregated data at the session level, and each entry is the difference between the average effort in the row position and the average effort in the column position. Significance levels refer to the Mann-Whitney test. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

An alternative way to investigate whether subjects in the high experience group make different choices due to their frequent random assignment to rare node types is to conduct a within subject analysis. Limiting our analysis to the high experience group of subjects previously defined, we can compare their choice of effort in the last round they have been assigned to a given node type to the $x$th time they have been assigned to that node type: a difference would indicate that the randomly determined degree of experience has an effect on the results because there is a non-negligible level of learning after the $x$th time a subject has been assigned to a given node type, where $x$ is equal to the total amount of times subjects in the low experience group have been assigned to that node type. Table 8 conducts this exercise for node types $H$ in $\mathbf{g}^{\otimes}$, $F$ in $\mathbf{g}^{15}$, and $F$ and $C$ in $\mathbf{g}^{21}$. There is no statistically significant difference in the choice of effort of subjects in the high experience group between the $x$th and the last time they have been assigned to any of the node types. Again, this suggests that the 40 rounds duration of the experiment is sufficient to guarantee that every subjects has enough experience to learn how to choose efforts even in rare node types.

## 5.2 Network visualization

In section 4.2 we analyzed the data by implicitly assuming that subjects consider as equivalent the different positions in each network that we have labelled to be the same node type in figure 1. Dessi et al. [2014] show that individuals utilize heuristics to memorize and recall information about social networks presented in graphical form, and these heuristics lead to biases in the recollection of aggregate and detailed aspects of the network structure. While it seems intuitive that individuals would consider positions labelled as the same node type as equivalent, this section presents an overview of appendix B where we conduct robustness checks to ensure that this assumption is validated in the data.

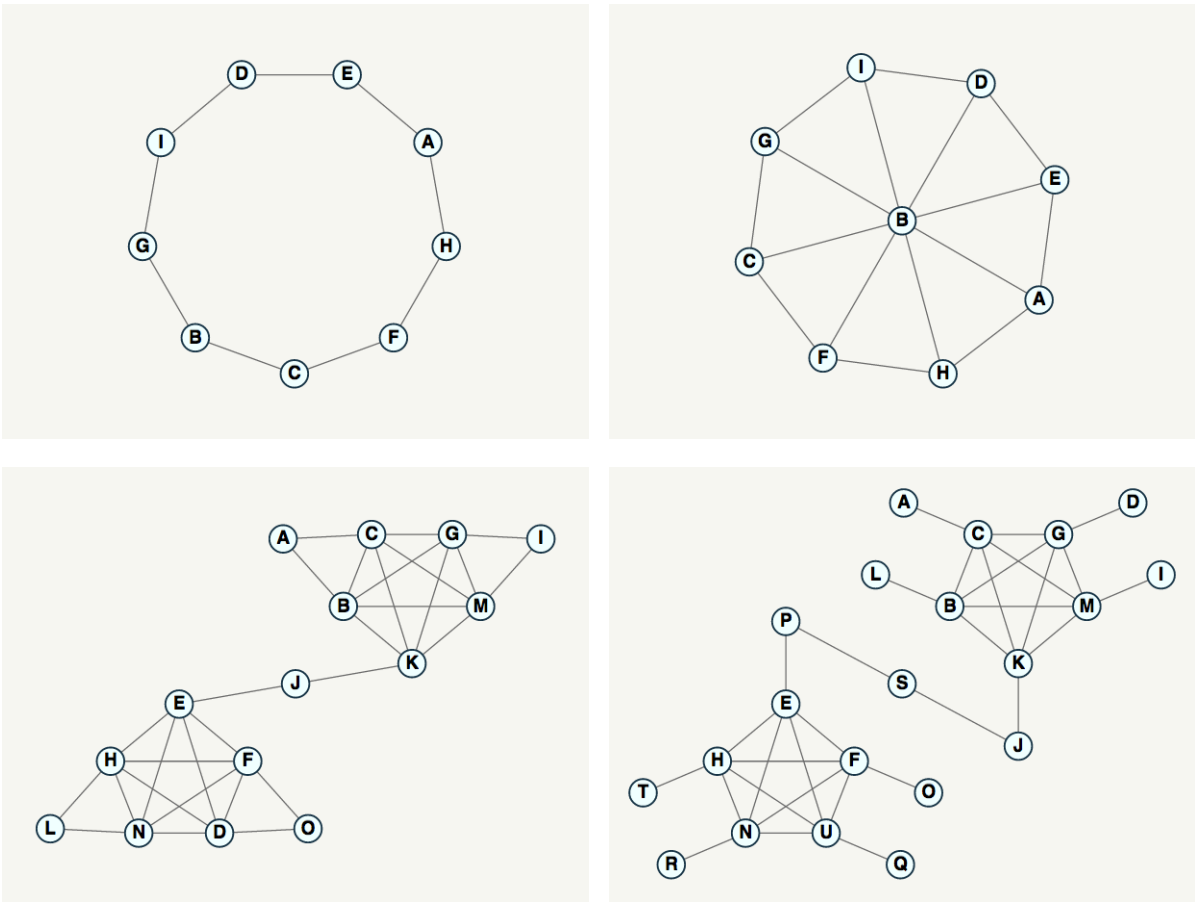



Figure 6: Network visualizations as seen by subjects in different treatments. Top: $\mathbf{g}^{\bigcirc}$ (left) and $\mathbf{g}^{\otimes}$ (right) treatments. Bottom: $\mathbf{g}^{15}$ (left) and $\mathbf{g}^{21}$ (right) treatments.

In the experimental interface, the network figure was always displayed in the top left part of the screen. Figure 6 shows how each network was actually displayed to the subjects, which is different from the graphical arrangement used in figure 1 for expository purposes in this paper. Tables 9, 10, 11 and 12 in appendix B are the equivalent of table 4 in section 4.2, but disaggregated for each node type for all the networks. For instance, table 9 shows the average difference in effort levels of subjects across the different positions which are labelled as node types $H$ and $B$ in the main analysis of the paper. A significant difference in effort levels between two positions would indicate that our assumption that these positions are equivalent may not hold. As it is clear from tables 9-12, across all the different pair of positions labelled as the same node type in the 4 networks, there are only 5 pairs of positions that have a statistically significant difference in average effort levels at the 10% level. This number of statistically significant findings is less than what

we would expect by chance given the large number of statistical tests, and therefore we conclude that our assumption that network positions labelled as the same node type are deemed equivalent by subjects is validated by the experimental data.

# 6 Conclusion

In this paper we have examined experimentally how social network structure mediates the tension between equilibrium play and the establishment of a more efficient norm by investigating a game of strategic complements played on different networks.

We find that subjects are able to coordinate on a collaborative norm when the network is simple and symmetric such as the circle, and, to a lesser extent, the wheel. In contrast, in the more complex networks of 15 and 21 nodes, on average subjects converge to the Nash equilibrium play on every node. Given the richness of the network structure and the size of the strategy space, this result provides a strong validation of the theoretical results in Ballester et al. [2006] and it shows how position in the network can be a key determinant of equilibrium play. These results are consistent with previous experimental work on public good games that shows that asymmetry across subjects leads to a decrease in the level of contributions. In future work, it would be interesting to explore systematically which structural features of the network leading to asymmetries across subjects have the largest detrimental impact on the coordination on efficient outcomes.

We also provide evidence that subjects base their decisions on the local, rather than the global, network structure. If we limit the analysis to the set of nodes in the 21 node network whose local features are not highly correlated with their overall position in the network, degree is a significant predictor of subjects' play, but Bonacich centrality, which determines equilibrium play, is not. In future work, it would be interesting to investigate whether focusing on local network features is a common way for individuals to cope with the challenging task of processing complex network information, independently of the game they are playing. This finding may also motivate further theoretical work: in many games played on networks it is challenging to solve for the equilibrium because strategies relate to the overall network structure in intricate ways, but this relation may simplify significantly if it is limited to the local network.

Methodologically, this paper shows the potential of UbiquityLab: a novel platform we developed to conduct online experiments with real-time interactions among participants. UbiquityLab is designed to be able to conduct online essentially any type of experiment that is currently run in physical labs. Moreover, as the experiment in this paper demonstrates, it rigorously incorporates features that replicate the state-of-the-art methodological practice in lab experiments with minimal adaptations which accommodate the specific requirements of the online set-up. The feasibility of online experiments with real-time interactions opens up exciting opportunities for experimental research including the possibility of creating a subject pool and running experimental sessions with a number

of participants which is orders of magnitude larger than current lab-based experiments, and the potential to run experiments across cultures with participants located anywhere in the world.

# References


C. Ballester, A. Calvó-Armengol, and Y. Zenou. Who's who in networks. Wanted: the key player. *Econometrica*, 74(5):1403–1417, 2006.

Y. Bramoullé, R. Kranton, and M. D'Amours. Strategic interaction and networks. *The American Economic Review*, 104(3):898–930, 2014.

T. Casciaro. Seeing things clearly: Social structure, personality, and accuracy in social network perception. *Social Networks*, 20:331–351, 1998.

A. Cassar. Coordination and cooperation in local, random and small world networks: Experimental evidence. *Games and Economic Behavior*, 58(2):209–230, 2007.

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

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

S. Choi, D. Gale, S. Kariv, and T. Palfrey. Network architecture, salience and coordination. *Games and Economic Behavior*, 73(1):76–90, 2011.

S. Choi, A. Galeotti, and S. Goyal. Trading in networks: Theory and experiments. 2014.

R. Dessi, E. Gallo, and S. Goyal. Network cognition. Cambridge-INET Working Paper Series 1413, July 2014.

D. M. Gale and S. Kariv. Trading in networks: A normal form game experiment. *American Economic Journal: Microeconomics*, pages 114–132, 2009.

A. Galeotti, S. Goyal, M. O. Jackson, F. Vega-Redondo, and L. Yariv. Network games. *Review of Economic Studies*, 77(1):218–224, 2010.

E. Gallo. Communication networks in markets. August 2014. Cambridge Working Papers in Economics 1431.

E. Gallo and C. Yan. The effects of reputational and social knowledge on cooperation. October 2014.

S. Goyal. *Connections: An Introduction to the Economics of Networks*. Princeton University Press, 2007.

C. Gracia-Lázaro, A. Ferrer, G. Ruiz, A. Tarancón, J. A. Cuesta, A. Sánchez, and Y. Moreno. Heterogeneous networks do not promote cooperation when humans play a prisoners dilemma. *Proceedings of the National Academy of Sciences*, 109(32):12922–12926, 2012.

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

M. O. Jackson. *Social and Economic Networks*. Princeton University Press, 2008.

M. Kearns, S. Suri, and N. Montfort. An experimental study of the coloring problem on human subject networks. *Science*, 313(5788):824–827, 2006.

M. Kearns, S. Judd, J. Tan, and J. Wortman. Behavioral experiments on biased voting in networks. *Proceedings of the National Academy of Sciences*, 106(5):1347–1352, 2009.

D. Krackhardt. Cognitive social structures. *Social Networks*, 9:109–134, 1987.

E. Kumbasar, K. A. Romney, and W. H. Batchelder. Systematic biases in social perception. *American Journal of Sociology*, 100(2):477–505, 1994.

S. K. Laury and C. A. Holt. Voluntary provision of public goods: Experimental results with interior Nash equilibria. *Handbook of Experimental Economics Results*, 1:792–801, 2008.

J. O. Ledyard. Public goods: A survey of experimental research. In J. Kagel and A. Roth, editors, *Handbook of Experimental Economics*. Princeton University Press, 1995.

E. Ostrom. *Governing the commons: The evolution of institutions for collective action*. Cambridge University Press, 1990.

A. Rapoport and R. Suleiman. Incremental contribution in step-level public goods games with asymmetric players. *Organizational Behavior and Human Decision Processes*, 55 (2):171–194, 1993.

S. Rosenkranz and U. Weitzel. Network structure and strategic investments: An experimental analysis. *Games and Economic Behavior*, 75(2):898–920, 2012.

H. P. Young. The evolution of social norms. *Annual Review of Economics* forthcoming, 2014.

# A UbiquityLab

*UbiquityLab* is a novel platform for creating and executing interactive experiments on the Internet. More broadly, the platform aims to promote online experiments for behavioral research in the social sciences by simplifying their creation and execution. The current version focuses on game-theoretic experiments. This section explains the functionalities and the workflow we use to conduct the experiments in this paper, and, where appropriate, it highlights additional available features.

In general, the platform accommodates two types of interaction. The first type is *real-time* interaction: subjects continuously act and respond to actions in a matter of seconds, and the whole experiment usually lasts no more than a few hours. For example, this experiment utilizes this type of interaction as subjects choose efforts and receive the results of their choices within 20 seconds over 40 consecutive rounds. The second type is *extended* interaction: subjects engage with each other irregularly and often in sequence, and the whole experiment may last days, weeks or even months. The current version emphasizes the support for real-time interactive experiments.

Computationally speaking, the platform can handle both synchronous and asynchronous processing of inputs from subjects. In the synchronous mode, a group of subjects play a $k$-stage ($k \geq 1$) game for $n$ rounds ($n \geq 1$). In each stage, some or all of the subjects in the group submit their actions before receiving feedback and proceeding to the next stage or round together. The stage-by-stage structure ensures that the actions of the current stage are processed at the same time before the start of the next stage. The experiment in this paper correspond to the $k = 1$ and $n = 40$ case.

In the asynchronous mode, the platform does not enforce the sequentiality and the processing of actions by stages/rounds: actions are processed individually as they arrive during a period, and an experiment may last for $n$ periods ($n \geq 1$). Here, a period represents a longer length of time than a short interval as in a stage/round in the synchronous mode. Unlike the synchronous mode, asynchronous processing does not need to wait for a batch of actions to complete before moving on to handle new ones. In addition, subjects may make different numbers of actions during a period. For example, market trading games may adopt asynchronous processing, as bids and asks are matched whenever they become available and subjects may trade different numbers of times in a period. In this mode, we can specify the maximum/minimum frequency of actions per subject per period as well as the duration of each period. Despite these underlying conceptual differences, the user experience is similar in both modes as participants make choices using a single graphical page on which new information updates automatically without clicking across multiple pages of text.

UbiquityLab operates a tested workflow that automates most phases of an experimental session. Before launching an experiment, we post a qualification HIT on AMT to recruit subjects using a simple survey with a fixed fee payment. We notify qualified subjects who have agreed to be willing to participate in the experiment about the date

and time of the sessions. Subjects enter sessions via an experiment HIT on AMT by clicking the embedded link in the HIT description. UbiquityLab can prevent multiple participation in the same or in a related experiment by checking the IP of participants and their AMT identification number. Notice that the platform supports running multiple experimental sessions in parallel within what we call a *meta-session* which corresponds to an experiment HIT. In this project, we ran up to 5 experimental sessions within a meta-session. In total, we ran 11 meta-sessions to finish all the treatments.

A meta-session starts with a virtual *Waiting Room*, where subjects remain active on a web page and wait until a specified number of subjects arrive. Once enough subjects get to the page, the waiting room automatically updates itself to allow the subjects to click a link to proceed into the *Instructions* of the experiment.

The instructions are divided into multiple pages. Depending on the content and the length of a page, the platform sets a minimum amount of time that a subject must spend on each page. If a subject attempts to proceed sooner than the minimum required time, a pop-up dialog appears to remind her to spend more time on that page. In the experiment in this paper, 'Previous' and 'Next' links allow subjects to navigate between pages, and the minimum required time ranges from 20 to 50 seconds. Notice that the minimum time limit only applies when a subject visits the page for the first time *and* she has not stayed for the required amount of time. In other words, once she stays beyond the minimum time, she does not need to wait if she visits the same page again.

After finishing the instructions, subjects must pass a *Quiz* before participating in the actual experiment. To pass the quiz, subjects usually need to answer all the questions correctly. The platform allows the experimenter to specify the number of attempts before failing a subject. In case of failure, the platform redirects the subject to a separate page and notifies the end of her participation. Furthermore, in order to ensure that subjects genuinely understand the instructions, the platform can randomize the details and the order of questions in each retake of the quiz. In the experiment in this paper, we allow up to three attempts to answer three questions regarding payoff calculation, and we randomize the specific values in each question on each retry.

Following the quiz, subjects enter the *Trial* phase where they play a trial version of the game to familiarize themselves with the user interface and understand the actual game. Typically, during the trial subjects play against a computer algorithm independently from each other. In the experiment in this paper, subjects play a 3-round trial game on a 5-node network where the efforts of neighbors were randomly generated in each round.

Once enough subjects complete the trial and arrive on the *Game* page, the actual experimental game starts automatically. The platform has different ways to ensure that the experimental game continues even if a subject exits playing due to a network error or a computer malfunction. Depending on the type of game being played, the computer can either notify the remaining players about the dropout before allowing them to continue the game without that subject, or replace the subject with a program player that plays

according to a pre-specified algorithm. The platform also supports random termination such that, after a specified number of rounds, the game terminates with a configurable probability in each subsequent round. In the experiment in this paper, the single-stage game repeats for 40 rounds with a deterministic termination, and a dropout subject is replaced by a computer algorithm that plays Nash. The data collected in a session involving dropout(s) is excluded from our analysis.

Finally, the experimenter may configure the platform to forward the subjects to a post-experimental *Survey* hosted on Qualtrics. After the questionnaire, subjects are automatically redirected to a page on the platform that summarizes the information about their earnings and provides them the code to input in AMT for payment. In the experiment in this paper, subjects complete a survey on Qualtrics that collects socio-demographic information and risk attitudes using the Holt and Laury [2002] test, and submit the payment confirmation code via the HIT page on AMT for the processing of payments using the Amazon Payment service.

UbiquityLab also offers two additional functionalities which have not been used in the experiment in this paper: the *Tour* and the *Chat*. The tour is a visual walkthrough of different components of the experimental interface, and it would typically follow the instructions. The chat offers the experimenter the capability of communicating with subjects during instructions and tour on either one-to-one or one-to-many basis. If activated, the chat is available on every page during those two phases.

The platform outputs data that allows the experimenter to monitor the progress of an ongoing session. The data includes metrics such as the number of subjects in each phase, the number of subjects who fail the quiz, the progress of the experimental game, etc. The current version presents the data via a self-updating log file with color-coding to differentiate the various types of information. In the next version, we are planning to have a more user-friendly web page for monitoring purposes.

Implementation wise, UbiquityLab consists of both server-side and client-side components. The server-side services are created using the Python programming language, while the client-side libraries are written in JavaScript. To develop an experimental game, the experimenter creates several minimal Python modules that plug into the servers using an Application Programming Interface (API). Each module specifies one particular aspect of the game such as initialization, per-stage processing, data recording, payment calculation, etc. A set of these modules defines a particular game. Thanks to this modular design, the platform offers the convenience and the freedom to implement highly customized game logic, while abstracting away many low-level intricacies. Furthermore, the experimenter is free to create any experimental interface that communicates with the servers via a provided messaging protocol. To ease the development effort, the platform offers a suite of common client-side functionalities such as timers, inactivity detection, input tracking, etc.

The objective of UbiquityLab is to allow researchers to exploit the potential of online

experiments without compromising on methodological rigour. Online experiments with UbiquityLab give researchers the freedom to conduct experiments anywhere and at anytime, which brings several advantages. First, it allows the creation of a diverse subject pool that is orders of magnitudes larger than current lab-based ones: this means more diversity in the subject pool as well as the capability of targeting the experiment to particular groups of interest. Second, it allows a large number of subjects to participate to the same experimental session. Third, it allows subjects located far away geographically to engage in the same experimental session opening the possibility of, e.g., cross-cultural studies. Fourth, there are several operational advantages including the possibility to run the experiments at any time of the year, lower costs, and a fast turnaround from design to implementation. Finally, UbiquityLab creates a controlled environment that is close to what can be achieved in the lab by monitoring and checking that subjects are actively engaged in the experimental task.

# B Further analysis

This appendix contains some further analysis of the experimental data. Tables 9-12 display the results of the analysis described in section 5.2 in the main text, which validate our implicit assumption that the specific network visualizations used in the experiment do not affect subjects' play. Figures 7-9 give further information on the evolution of efforts and payoffs for each position type in each treatment.

Table 9: Differences in average effort across different positions for node types $H$ and $B$ for $\mathbf{g}^{21}$.

| | | | $\mathbf{g}^{21}$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | **H** | | | | | | | **B** |
| | | | 2 | 3 | 4 | 14 | 15 | 16 | 17 | 5 |
| $\mathbf{g}^{21}$ | **H** | 1 | -3.4 | -0.6 | -0.3 | 2.2 | -4.5 | -1.3 | -3.9 | |
| | | 2 | | 2.7 | 3.0 | 5.6 | -1.1 | 2.1 | -0.5 | |
| | | 3 | | | 0.3 | 2.9 | -3.9 | -0.7 | -3.3 | |
| | | 4 | | | | 2.6 | -4.2 | -1.0 | -3.6 | |
| | | 14 | | | | | −6.7 | -3.5 | −6.1 | |
| | | 15 | | | | | | 3.2 | 0.6 | |
| | | 16 | | | | | | | -2.6 | |
| | **B** | 13 | | | | | | | | 2.7 |

The analysis is based on aggregated data at the session level, and each entry is the difference between the average effort in the row position and the average effort in the column position. Significance levels refer to the Mann-Whitney tests.$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 10: Differences in average effort across different positions for node types $P$ and $C$ for $\mathbf{g}^{15}$ and $\mathbf{g}^{21}$.

| | | | $\mathbf{g}^{15}$ | | | $\mathbf{g}^{21}$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | **P** | | | **P** | | | | | | | **C** |
| | | | 7 | 14 | 15 | 7 | 8 | 9 | 18 | 19 | 20 | 21 | 10 |
| $\mathbf{g}^{15}$ | **P** | 6 | 2.2 | 1.5 | 1.6 | | | | | | | | |
| | | 7 | | -0.7 | -0.7 | | | | | | | | |
| | | 14 | | | 0.0 | | | | | | | | |
| $\mathbf{g}^{21}$ | **P** | 6 | | | | -1.9 | $-6.1^{*}$ | 1.1 | -1.3 | -2.3 | -0.9 | -2.1 | |
| | | 7 | | | | | -4.2 | 3.1 | 0.6 | -0.3 | 1.1 | -0.2 | |
| | | 8 | | | | | | $7.2^{*}$ | 4.8 | 3.8 | $5.2^{*}$ | 4.0 | |
| | | 9 | | | | | | | -2.4 | -3.4 | -2.0 | -3.3 | |
| | | 18 | | | | | | | | -1.0 | 0.4 | -0.8 | |
| | | 19 | | | | | | | | | 1.4 | 0.1 | |
| | | 20 | | | | | | | | | | -1.3 | |
| | **C** | 11 | | | | | | | | | | | -1.5 |

The analysis is based on aggregated data at the session level, and each entry is the difference between the average effort in the row position and the average effort in the column position. Significance levels refer to the Mann-Whitney tests. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 11: Differences in average effort across different visual positions for node type $H$ for $\mathbf{g}^{15}$.

| | | $\mathbf{g}^{15}$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 2 | 3 | 4 | 5 | 9 | 10 | 11 | 12 | 13 |
| $\mathbf{g}^{15}$ | 1 | 0.5 | 3.0 | -0.4 | 0.4 | 1.1 | 1.0 | 3.1 | 2.2 | 0.5 |
| | 2 | | 2.5 | -0.9 | -0.1 | 0.6 | 0.5 | 2.6 | 1.7 | 0.0 |
| | 3 | | | -3.5 | -2.6 | -1.9 | -2.1 | 0.1 | -0.8 | -2.5 |
| | 4 | | | | 0.8 | 1.5 | 1.4 | 3.6 | 2.6 | 0.9 |
| | 5 | | | | | 0.7 | 0.6 | 2.7 | 1.8 | 0.1 |
| | 9 | | | | | | -0.1 | 2.0 | 1.0 | -0.6 |
| | 10 | | | | | | | 2.2 | 1.2 | -0.5 |
| | 11 | | | | | | | | -0.9 | -2.6 |
| | 12 | | | | | | | | | -1.7 |

The analysis is based on aggregated data at the session level, and each entry is the difference between the average effort in the row position and the average effort in the column position. Significance levels refer to the Mann-Whitney tests. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 12: Differences in average effort across different visual positions for nodes in $\mathbf{g}^{\bigcirc}$ and node type $P$ in $\mathbf{g}^{\otimes}$.

| | | $\mathbf{g}^{\bigcirc}$ | | | | | | | | $\mathbf{g}^{\otimes}$ | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| $\mathbf{g}^{\bigcirc}$ | 1 | 0.8 | -1.1 | 0.8 | 0.9 | 0.4 | 0.4 | -0.5 | -0.2 | | | | | | | |
| | 2 | | -1.9 | -0.1 | 0.1 | -0.4 | -0.4 | -1.3 | -1.0 | | | | | | | |
| | 3 | | | 1.8 | 1.9* | 1.5 | 1.5 | 0.6 | 0.9 | | | | | | | |
| | 4 | | | | 0.1 | -0.3 | -0.3 | -1.2 | -0.9 | | | | | | | |
| | 5 | | | | | -0.5 | -0.4 | -1.3 | -1.0 | | | | | | | |
| | 6 | | | | | | 0.0 | -0.9 | -0.6 | | | | | | | |
| | 7 | | | | | | | -0.9 | -0.6 | | | | | | | |
| | 8 | | | | | | | | 0.3 | | | | | | | |
| $\mathbf{g}^{\otimes}$ | 2 | | | | | | | | | 1.0 | 0.1 | -0.1 | 1.2 | 1.7 | 1.1 | 0.4 |
| | 3 | | | | | | | | | | -0.9 | -1.1 | 0.1 | 0.6 | 0.0 | -0.6 |
| | 4 | | | | | | | | | | | -0.2 | 1.0 | 1.5 | 0.9 | 0.2 |
| | 5 | | | | | | | | | | | | 1.3 | 1.8 | 1.2 | 0.5 |
| | 6 | | | | | | | | | | | | | 0.5 | -0.1 | -0.8 |
| | 7 | | | | | | | | | | | | | | -0.6 | -1.3 |
| | 8 | | | | | | | | | | | | | | | -0.7 |

The analysis is based on aggregated data at the session level, and each entry is the difference between the average effort in the row position and the average effort in the column position. Significance levels refer to the Mann-Whitney tests. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

.

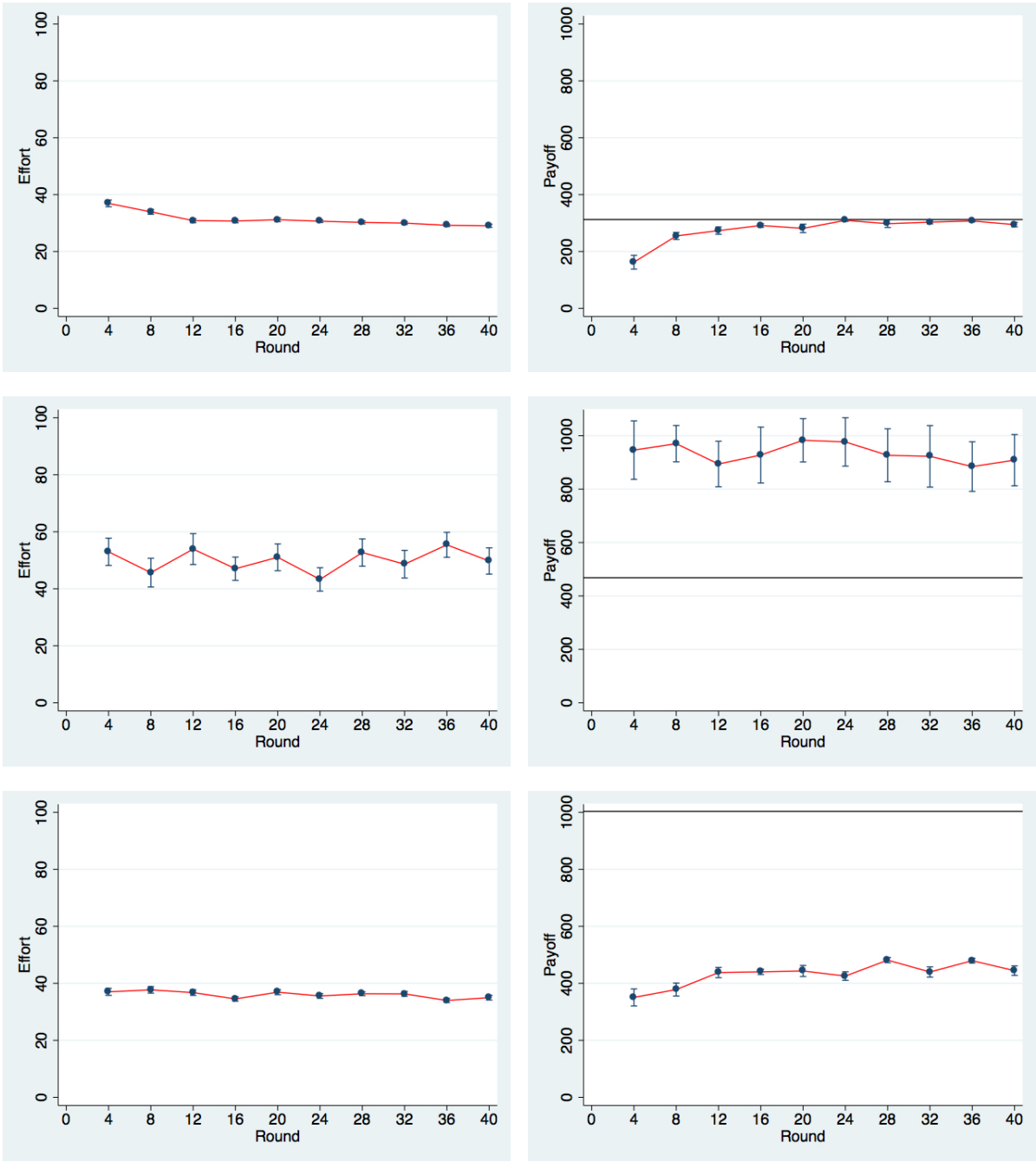



Figure 7: Evolution of efforts and payoffs (aggregated every 4 rounds) in networks $\mathbf{g}^{\bigcirc}$ and $\mathbf{g}^{\otimes}$. Top: nodes in $\mathbf{g}^{\bigcirc}$. Middle: node type $H$ in $\mathbf{g}^{\otimes}$. Bottom: node type $P$ in $\mathbf{g}^{\otimes}$. Lines indicate predicted Nash (black) and mean (red) effort or payoff levels. Error bars indicate SEM.

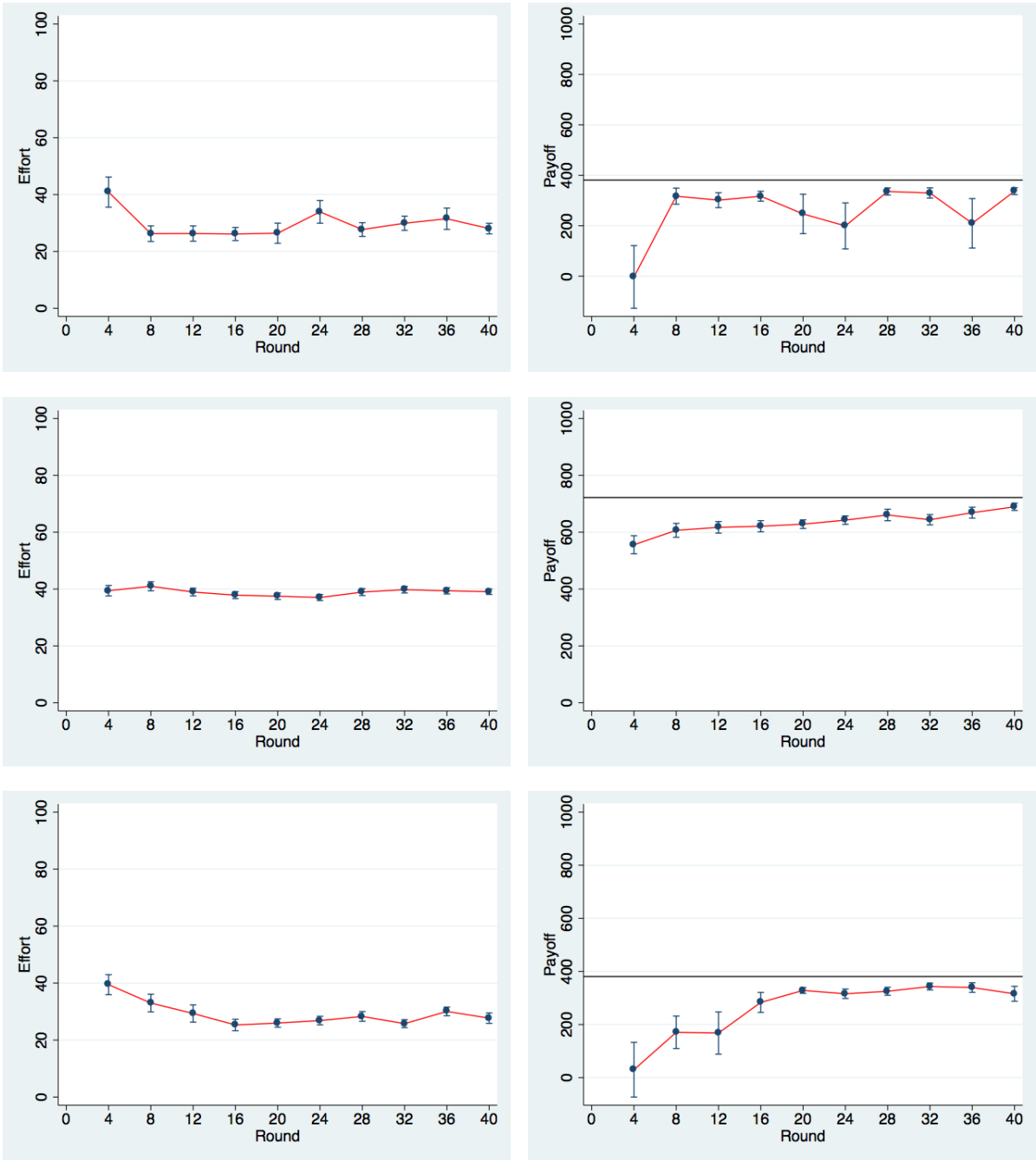



Figure 8: Evolution of efforts and payoffs (aggregated every 4 rounds) in network $\mathbf{g}^{15}$. Top: node type $F$. Middle: node type $H$. Bottom: node type $P$. Lines indicate predicted Nash (black) and mean (red) effort or payoff levels. Error bars indicate SEM.

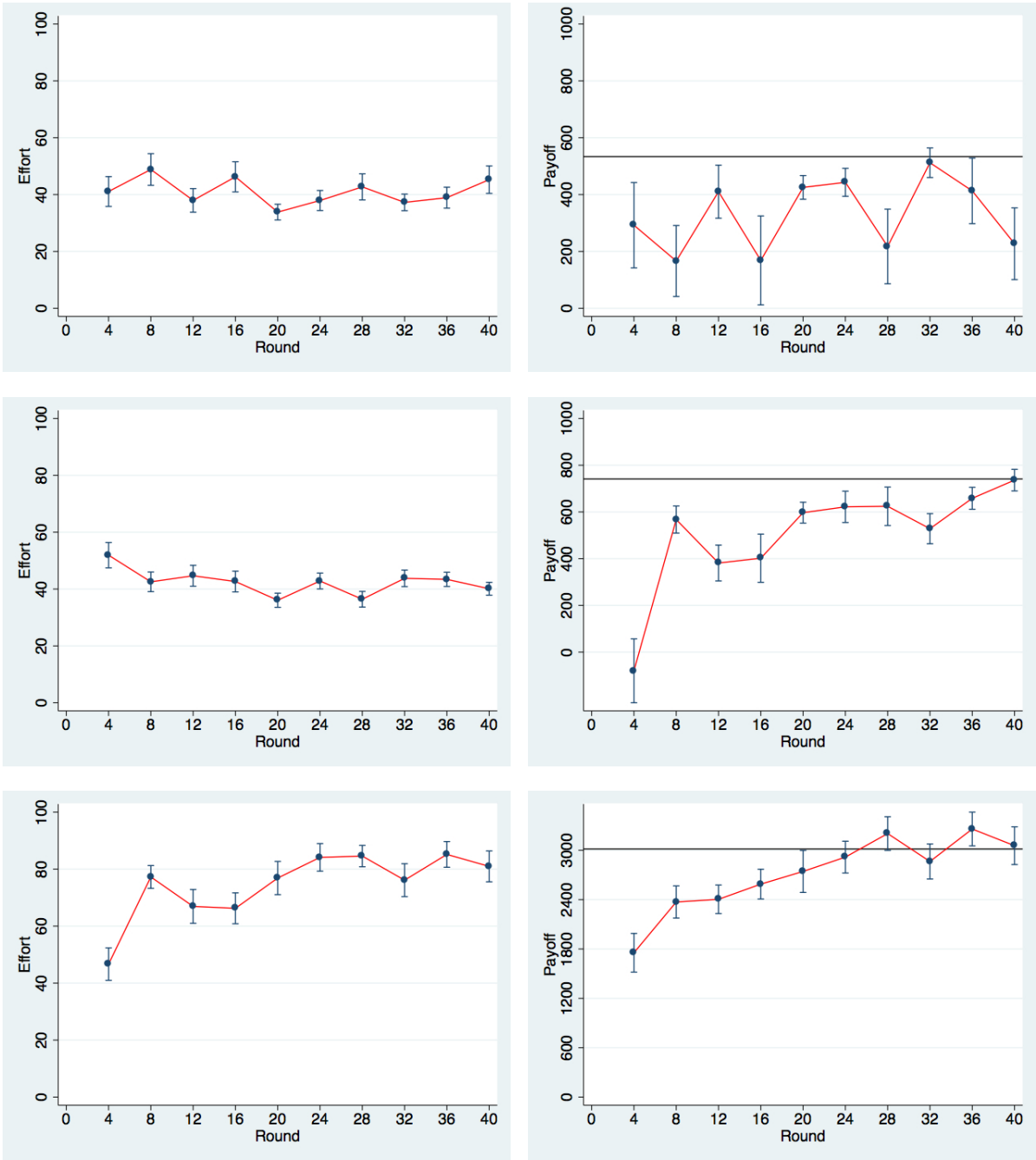



Figure 9: Evolution of efforts and payoffs (aggregated every 4 rounds) in network $\mathbf{g}^{21}$. Top: node type $F$. Middle: node type $C$. Bottom: node type $B$. Lines indicate predicted Nash (black) and mean (red) effort or payoff levels. Error bars indicate SEM.

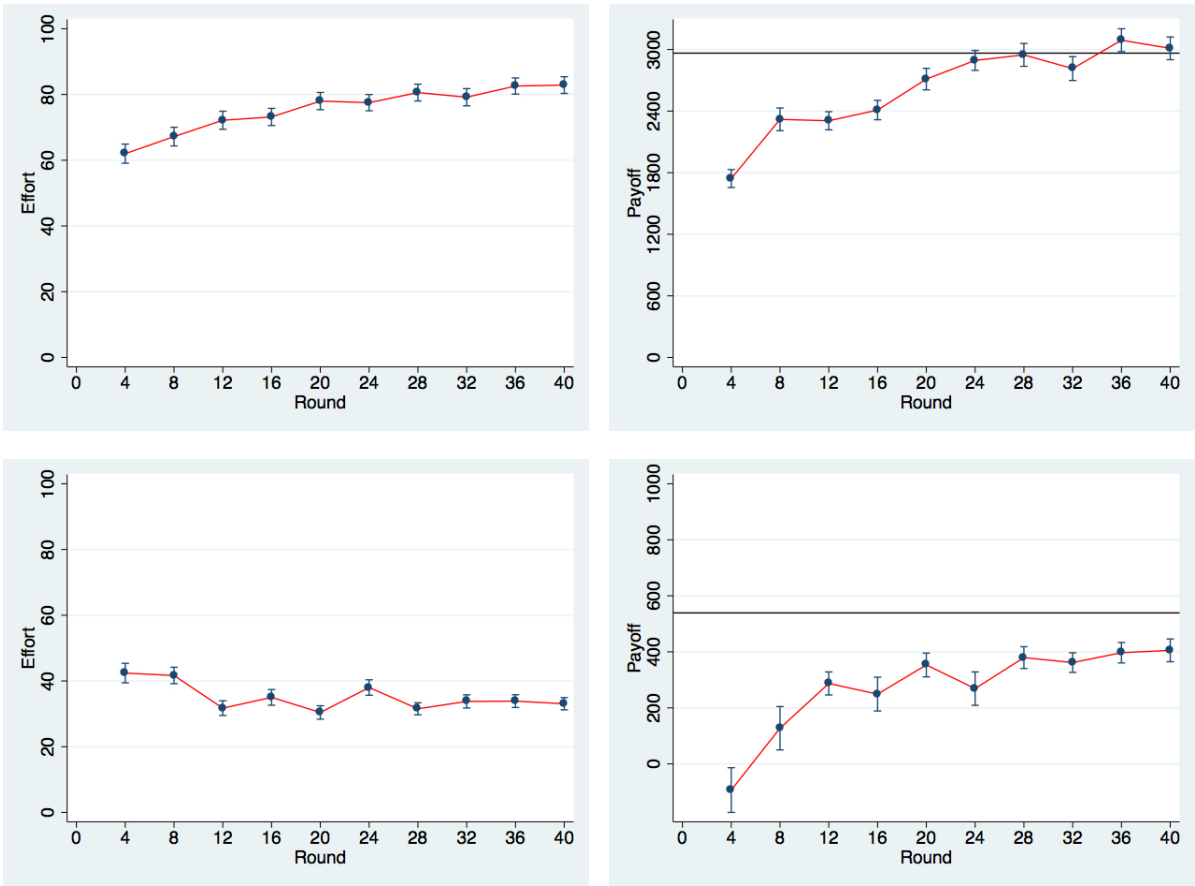



Figure 9: Evolution of efforts and payoffs (aggregated every 4 rounds) in network $\mathbf{g}^{21}$. Top: node type $H$. Bottom: node type $P$. Lines indicate predicted Nash (black) and mean (red) effort or payoff levels. Error bars indicate SEM.

# C Experimental material

## C.1 Instructions and Quiz

This section shows the screenshots of the instructions and the quiz before the experiment.

## General rules 1/5

The aim of this Experiment is to study how individuals make decisions in certain contexts. You will be asked to make decisions that will affect the amount of points you earn and the amount of points other Turkers earn. It is important for you to know that **your decisions will remain completely confidential**. Each person will be assigned fictitious initials: **we will always use your fictitious initials and never your Turker ID or any other information that might allow other participants to identify you.**

It is important that **you read these instructions carefully**. Note that each participant is shown exactly the same instructions. At the end of these instructions you will be asked to take a **Quiz** to ensure that you understand the instructions. If you fail to solve the Quiz, you will not be able to participate in the Experiment.

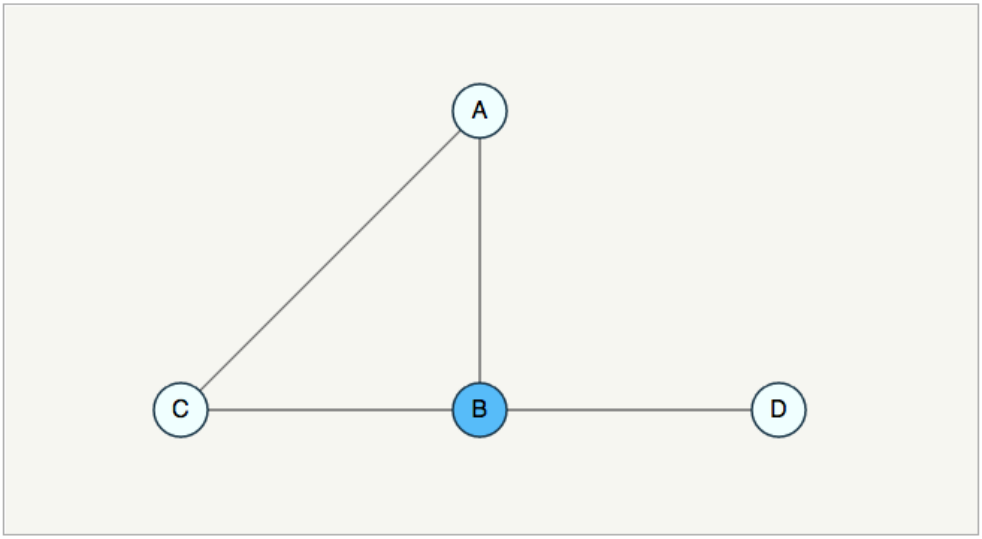
Please click the **'Next'** link below if you would like to continue to the description of the experiment.

# General setup 2/5

You and the other Turkers participating in this Experiment are placed on a "network" and the participants who are directly connected to you in this network are your "network neighbors." It is important to understand that the points you win are determined only by how you interact with these network neighbors (i.e. participants directly linked to you in the network).

The Experiment consists of 40 rounds. The network is the same for all 40 rounds and it will be displayed on the left side of the screen throughout the Experiment. Each position in the network is labelled by a capital letter. In each round you will be randomly assigned to one of the 9 positions in the network. In each round, you will be informed of your position (letter) and the node representing your position will be coloured in **blue**.

For example, in the network below you are in position B. In a network, a "link" is represented by a line between two positions. The position B you are assigned to has 3 links with positions A, C and D. The Turkers assigned to positions A, C and D are your neighbors.

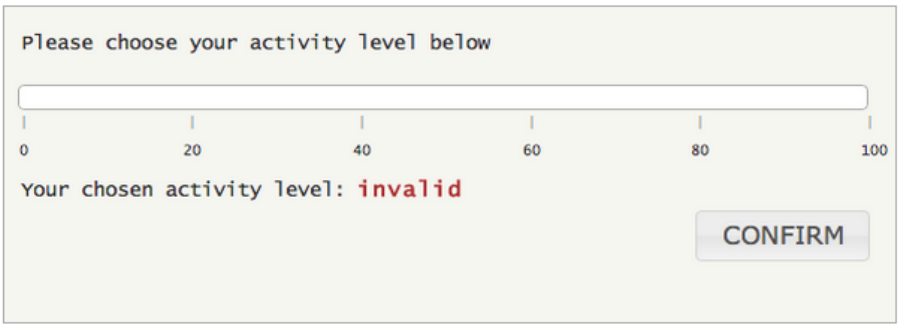



Please click the **"Next"** link below to continue to the description of how your own and your neighbors' decisions will affect the number of points you win.

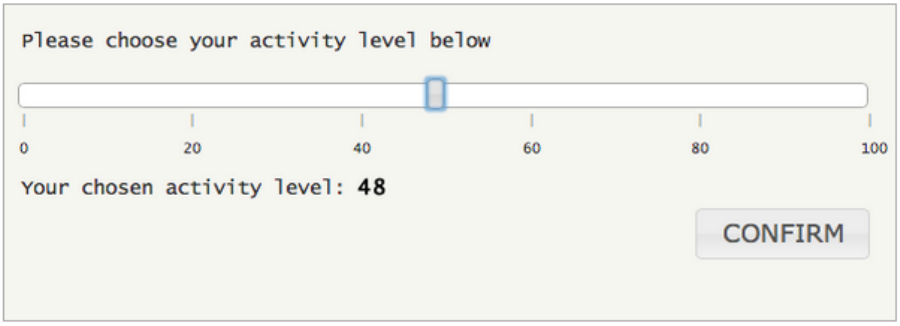
# Choice of activity level 3/5

At each round, knowing your position, you will be asked to make a choice on your **level of activity**. The minimum activity level is 0, the maximum is 100 and you can pick any level in between.

You pick your activity level by pointing the mouse and clicking on the scale, as shown below. Once you have clicked on the scale, a handle bar will appear. You can select the exact choice of your activity level either by clicking on the scale using your mouse or by using the left/right arrow buttons on your keyboard. Notice that if you do not select a level of activity then your chosen activity level is "invalid" and **you will be deducted 2,000 points** for not participating in the round.




You have a maximum of 20 seconds to select your level of activity. After the first 5 seconds, the slider and a "Confirm" button will appear. You then have 15 seconds to select and confirm your activity level. For example, in the figure below, a level of activity equal to 48 is selected. If you do not click the "Confirm" button by the end of 20 seconds, you will be considered inactive in this round and you will be deducted 2,000 points. The round will end when the 20 seconds pass or whenever all Turkers have confirmed their level of activity, whichever comes earlier.



You earn $1 by successfully completing the Experiment and the Final Questionnaire, independently on your performance in the Experiment. Additionally, you can earn points during the Experiment and **these points will be converted into your bonus at the end of the Experiment**. The conversion procedure will be explained at the end of the instructions. The points you obtain depend in part from your activity level and in part from the activity level of your neighbours.

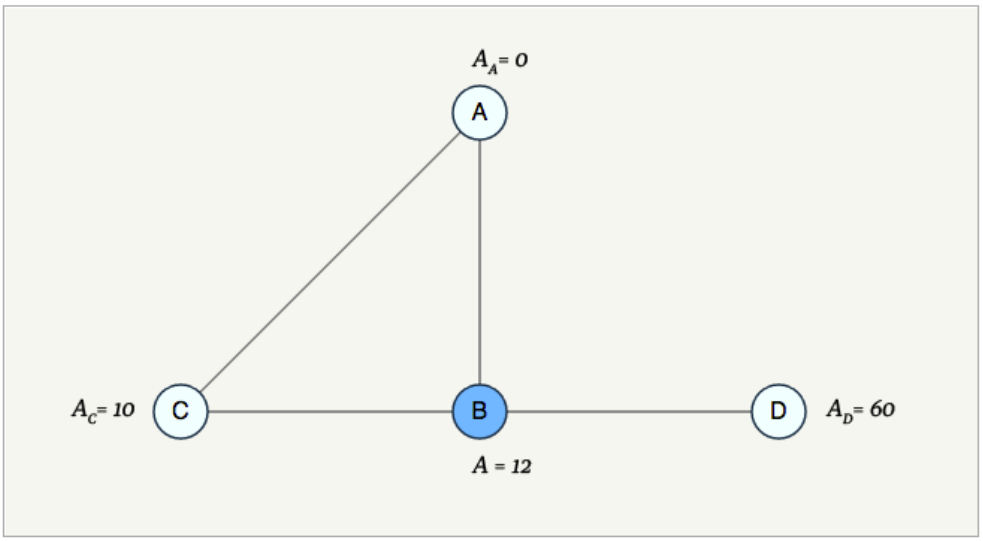
Please click the **"Next"** link below to continue to the description of how your own and others' decisions will affect the number of points you win.

# How to win points 4/5

The number of points you win in each round depends on **your own and your neighbors' activity levels**. Suppose that you pick a level of activity $A$, the number of points you win is computed according to the following formula:

$$\text{Points Won} = 20A - \frac{1}{2}A^2 + \frac{1}{10}A \times (\text{sum of your neighbors' activity levels})$$

Now imagine that you are in position B in the network below. You have 3 neighbors A, C and D.




Suppose that you pick a level of activity $A$ and that your neighbors select activity levels $A_A$, $A_C$ and $A_D$, then the number of total points you win is:

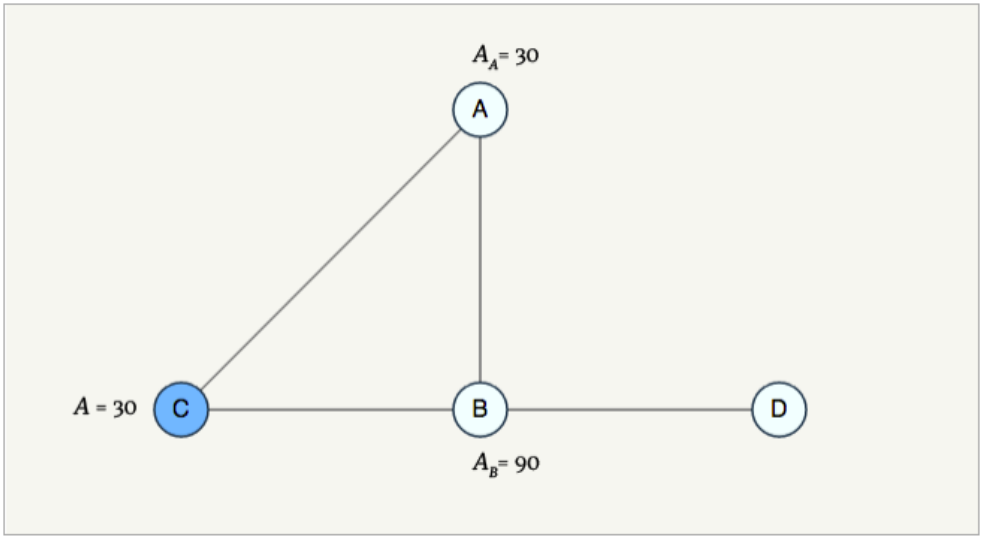
$$\text{Points Won} = 20A - \frac{1}{2}A^2 + \frac{1}{10}AA_A + \frac{1}{10}AA_C + \frac{1}{10}AA_D$$

For instance, if you select a level of activity $A = 12$ and your neighbors select activity levels $A_A = 0$, $A_C = 10$ and $A_D = 60$, then the number of total points you win is:

$$\text{Points Won} = 20 \times 12 - \frac{1}{2} \times 12^2 + \frac{1}{10} \times 12 \times 0 + \frac{1}{10} \times 12 \times 10 + \frac{1}{10} \times 12 \times 60$$
$$= 240 - 72 + 0 + 12 + 72 = 252$$

Let us go through another example to make sure you understand how you win points in the experiment.

Imagine that you are in position C in the network below. You have 2 neighbors A and B.




Suppose that you pick a level of activity $A = 30$ and that your neighbors select activity levels $A_A = 30$ and $A_B = 90$, then the number of total points you win is:

$$\text{Points Won} = 20A - \frac{1}{2}A^2 + \frac{1}{10}AA_A + \frac{1}{10}AA_B$$
$$= 20 \times 30 - \frac{1}{2} \times 30^2 + \frac{1}{10} \times 30 \times 30 + \frac{1}{10} \times 30 \times 90$$
$$= 600 - 450 + 90 + 270 = 510$$

At the end of the round you will see the following information:

- Your current network position
- Your current network neighbors
- Your chosen activity level
- Your neighbors' chosen activity levels
- Total points won in this round

Please click the **"Next"** link below to continue to the description of how the points will be converted into earnings.

# Your earnings 5/5

At the end of the experiment, 8 out of 40 rounds will be selected randomly to determine your earnings. The exchange rate from points (EP) to US dollars is:

$$20 \text{ EP} = 1 \text{ cent}$$

Let $V$ = number of EP won in the 8 randomly drawn rounds, then your total earnings will be equal to:

$$\text{Total earnings} = 100 + 1 \times \frac{V}{20} \text{ cents}$$

For example, if $V$ = 6,000 points then at the end of the experiment you will earn:

$$\begin{aligned} \text{Total earnings} &= 100 + 1 \times \frac{V}{20} \\ &= 100 + 1 \times \frac{6000}{20} \\ &= 100 + 300 = 400 \text{ cents} \end{aligned}$$

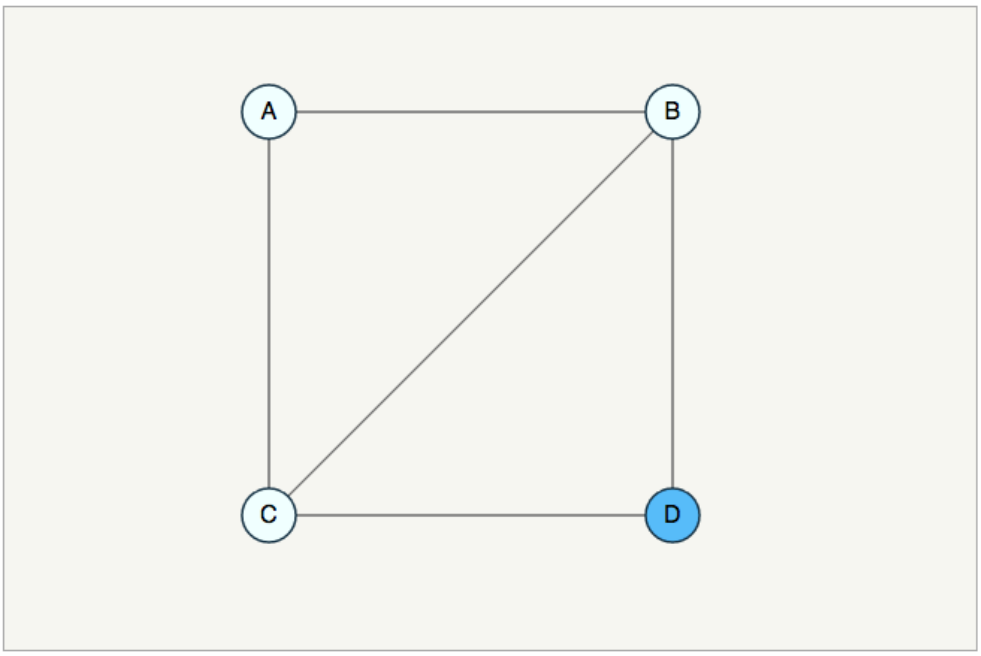
You can also win an additional Bonus payment in the Final Questionnaire. We will explain this when you are doing the Final Questionnaire.

Please click the **"Next"** link below to continue to a Quiz for testing your understanding of the experiment.

# Quiz

To make sure you have read and understood the instructions, you must answer the following questions correctly.

Assume you are the Blue node D in the network below. You have two neighbors B and C.




Given that your activity level is $A$ and your neighbors' activity levels are $A_B$ and $A_C$ respectively, the number of points that you earn is equal to:

$$\text{Points Won} = 4A - \frac{1}{2}A^2 + \frac{1}{3}AA_B + \frac{1}{3}AA_C$$

Please note: each answer box below will turn **green** when you enter a correct answer, and will turn **red** when you enter a wrong answer. You have to answer all the questions correctly before continuing.

1. If your activity level is $A$ = 8 and your neighbors' activity levels are $A_B$ = 0 and $A_C$ = 0, then what is the total number of points you win?

   Answer

2. If your activity level is $A$ = 0 and your neighbors' activity levels are $A_B$ = 61 and $A_C$ = 63 then what is the total number of points you win?

   Answer

3. If your activity level is $A$ = 30 and your neighbors' activity levels are $A_B$ = 0 and $A_C$ = 60 then what is the total number of points you win?

   Answer

After passing this quiz, you will play 3 trial rounds of the Experiment on a network with 5 positions. The purpose of the trial rounds is to get you familiar with the user interface of the Experiment and the amount of time available for you to select an activity level.

**These trial rounds have no impact on the number of points that you win**. In each trial round every Turker is randomly assigned to a position and the activity levels of the neighbors are randomly drawn between 0 and 100.

Please click the **"Next"** link below to start the trial rounds.

## C.2 Post-experimental Questionnaire

This section shows the screenshots of the post-experimental questionnaire after the game.

1. What is your gender?

Female ◯ Male ◯

2. What year were you born?

3. What is your nationality?

4. What is the highest level of education you have completed?

- ◯ Less than High School
- ◯ High School / GED
- ◯ Some College
- ◯ 2-year College Degree
- ◯ 4-year College Degree
- ◯ Masters Degree
- ◯ Doctoral Degree
- ◯ Professional Degree (JD, MD, etc)

5. Which occupational category best describes your employment? (U.S. Census, 40 Categories)

>>

6. Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?

- ◯ Most people can be trusted
- ◯ Need to be very careful

>>

**Bonus question!**
This question determines your potential additional bonus payment. We will randomly select 5% of all the Turkers participating in the Experiment and pay them the bonus that they win in The Decision Game below. Within 3 days we will pay the bonus to the winners and publish their Worker IDs at this URL: https://www.playmymodel.com/notice/decision-game/draw-results

**Decision Game**

You have ten Decisions to make. Each Decision is a choice between "Option A" and "Option B." You will have to make ten choices by selecting either Option A or Option B for each of the decisions.

In order to understand what the decisions are, please look at Decision 1 at the top. Option A pays 200 cents with 10% probability, and it pays 160 cents with 90% probability. Option B pays 385 cents with 10% probability, and it pays 10 cents with 90% probability. The other Decisions are similar, except that as you move down the table the chances of the higher payoff for each option increase. In fact, for Decision 10 at the bottom, each option pays the highest payoff for sure, so your choice here is between 200 cents or 385 cents. Note that in each option you always get paid something, what changes is the probability that you get paid a high or a low amount.

To summarize, you will make ten choices: for each decision row you will have to choose between Option A and Option B. You may choose A for some decision rows and B for other rows, and you may change your decisions and make them in any order.

Your bonus will be computed as follows. First, the computer will randomly pick one of the 10 Decisions. Second, we will look at the Option that you choose for that decision and you will get either the high or low amount with probabilities given by the chosen Option. For instance, if the computer randomly selected Decision 1 at the top and you chose Option B then you will win 385 cents with 10% probability and 10 cents with 90% probability. We will tell you the outcome of the Decision Game at the end of the survey. We will draw the winners within 3 days, and if you are one of the 5% of Turkers who are selected, you will receive the payment within 3 days.

| A | B |
|---|---|
| 10% chance of winning $2.00 and 90% chance of winning $1.60 | 10% chance of winning $3.85 and 90% chance of winning $0.10 |
| 20% chance of winning $2.00 and 80% chance of winning $1.60 | 20% chance of winning $3.85 and 80% chance of winning $0.10 |
| 30% chance of winning $2.00 and 70% chance of winning $1.60 | 30% chance of winning $3.85 and 70% chance of winning $0.10 |
| 40% chance of winning $2.00 and 60% chance of winning $1.60 | 40% chance of winning $3.85 and 60% chance of winning $0.10 |
| 50% chance of winning $2.00 and 50% chance of winning $1.60 | 50% chance of winning $3.85 and 50% chance of winning $0.10 |

| | |
|---|---|
| 60% chance of winning $2.00 and 40% chance of winning $1.60 | 60% chance of winning $3.85 and 40% chance of winning $0.10 |
| 70% chance of winning $2.00 and 30% chance of winning $1.60 | 70% chance of winning $3.85 and 30% chance of winning $0.10 |
| 80% chance of winning $2.00 and 20% chance of winning $1.60 | 80% chance of winning $3.85 and 20% chance of winning $0.10 |
| 90% chance of winning $2.00 and 10% chance of winning $1.60 | 90% chance of winning $3.85 and 10% chance of winning $0.10 |
| 100% chance of winning $2.00 and 0% chance of winning $1.60 | 100% chance of winning $3.85 and 0% chance of winning $0.10 |

>>

The computer has randomly selected Decision 9 out of the 10 possible decisions. For this decision, you chose Option A: 90% chance of winning $2.00 and 10% chance of winning $1.60.

The computer has randomly chosen the outcome of your decision: you have won $2. In case that you are one of the 5% of Turkers who are randomly selected to get a bonus, you will be paid a bonus of $2.

>>

Please briefly describe how you picked your activity level during the experiment.

(Optional) You are welcome to make any comment about the experiment and/or your experience participating in this HIT.

>>