UNIVERSITY OF CAMBRIDGE
Faculty of Economics

JANEWAY INSTITUTE CAMBRIDGE

CAMBRIDGE WORKING PAPERS IN ECONOMICS
JANEWAY INSTITUTE WORKING PAPERS

Learning in Canonical Networks

Syngjoo Choi
Seoul National University

Sanjeev Goyal
University of Cambridge

Frederic Moisan
Emlyon Business School, GATE

Yu Yang Tony To
University of Cambridge

Abstract

Subjects observe a private signal and then make an initial guess; they observe their neighbors' guesses and guess again, and so forth. We study learning dynamics in three networks: Erdös-Rényi, Stochastic Block (reflecting homophily) and Royal Family (that accommodates both highly connected celebrities and local interactions). We find that the Royal Family network is more likely to sustain incorrect consensus and that the Stochastic Block network is more likely to persist with diverse beliefs. These aggregate patterns are consistent with individuals following DeGroot updating rule.

Reference Details

2235	Cambridge Working Papers in Economics
2022/12	Janeway Institute Working Paper Series
Published	1 June 2022
Key Words	social learning, social networks, experimental social science, consensus
JEL-codes	C91, C92, D83, D85
Websites	www.econ.cam.ac.uk/cwpe www.janeway.econ.cam.ac.uk/working-papers

Learning in Canonical Networks

Syngjoo Choi* Sanjeev Goyal† Frederic Moisan‡ Yu Yang Tony To§

June 1, 2022

Abstract

Subjects observe a private signal and then make an initial guess; they observe their neighbors' guesses and guess again, and so forth. We study learning dynamics in three networks: Erdös-Rényi, Stochastic Block (reflecting homophily) and Royal Family (that accommodates both highly connected celebrities and local intearctions). We find that the Royal Family network is more likely to sustain incorrect consensus and that the Stochastic Block network is more likely to persist with diverse beliefs. These aggregate patterns are consistent with individuals following DeGroot updating rule.

Keywords: social learning, social networks, experimental social science, consensus

JEL Codes: C91, C92, D83, D85

*Seoul National University, syngjooc@snu.ac.kr
†University of Cambridge, sg472@cam.ac.uk
‡Emlyon Business School, GATE, moisan@em-lyon.com
§University of Cambridge, yytt2@cam.ac.uk
We thank Matthew Elliot, Edoardo Gallo, Ben Golub, Fulin Guo, Sam Jindani, Omer Tamuz, Leeat Yariv and participants at a number of seminars for useful comments. We thank Keynes Fund (University of Cambridge), Creative-Pioneering Researchers Program (Seoul National University) and C-BID (NYUAD) for financial support.

1 Introduction

> In these democratic days, any investigation into the trustworthiness and peculiarities of popular judgements is of interest. *Galton (1907), pages 450-451.*

More than a hundred years after Galton's discovery of the "wisdom of crowds" (Galton, 1907), as democratic politics became more common across the world, our collective opinions and beliefs matter for an ever-widening range of subjects. Pioneering work on the role of social networks was carried out by sociologists in the mid-twentieth century (Lazarsfeld and Merton, 1954, Katz and Lazarsfeld, 1966, Coleman et al., 1966). More recently, with the growing usage of social media, there has been renewed interest in the role of social networks in shaping opinion formation and behaviour. Existing studies have highlighted two features of real-world social networks: (i) deep inequalities in the number of connections where the average is small but the variance is very large, and (ii) network homophily — tendency of people with similar traits to form links with each other (Barabási and Albert, 1999, Newman, 2010, McPherson et al., 2001, Currarini et al., 2009). The theory of social learning shows that these network features have powerful effects on opinions and behaviour (Bala and Goyal (1998), Bala and Goyal (2001), DeMarzo et al. (2003), Mossel et al. (2014) and Golub and Jackson (2010)); for a survey of this research see Golub and Sadler (2016) and Goyal (ming). This paper aims to experimentally test these theoretical predictions in large canonical networks, i.e., networks that are rich and complex and that reflect inequality and homophily.

We consider a model taken from Gale and Kariv (2003) in which individuals receive noisy signals about the true state of the world and make a guess repeatedly over time. We consider a binary state setting with a binary guess where the optimal guess is to match the true state. Individuals also observe the guesses of their neighbours, which in principle allows information to flow across paths of the social network. We examine how the network shapes the long-run process of information dissemination.

We study learning in three networks: Erdös-Rényi (a baseline for connections among homogeneous individuals), Stochastic Block (reflecting network homophily) and Royal Family network (that accommodates 'influential individuals' along with local interactions). Figure 1 presents these three networks and Figures 2a and 2b present the learning dynamics under DeGroot updating (DeGroot, 1974): at any period t, an individual guesses the state that corresponds to the majority guess in her neighbourhood in the previous period $t-1$. We are led to three hypotheses: (i) individual behaviour converges; (ii) the presence of network homophily leads to the persistence of diverse opinions/guesses; (iii) the presence of influential individuals gives rise to incorrect consensus and sub-optimal behaviour. In real life, people are diverse in preferences, capacities for information processing, and decision-making rules. It is therefore unclear if these theoretical predictions will obtain in practice.

We conduct a laboratory experiment to test these predictions.[1] Our experiments yield three findings. First, learning occurs in all the networks so rapidly that most of the consensus level achieved happens early. Second, breakdown of consensus and persistence of diverse opinions is more likely in the Stochastic Block network as compared to the other two networks. Third, incorrect consensus is much more likely in the Royal Family network as compared to the other two networks. Finally, we show that the vast majority of individual guesses are consistent with DeGroot updating rule.

Related Literature. There is a large body of experimental research on opinion formation and behaviour. Early contributions include Choi et al. (2005), Mobius et al. (2015), Kearns et al. (2012). For a survey of the experimental research in economics see Choi et al. (2016), Breza (2016). Our paper is closest to two recent papers by Grimm and Mengel (2020) and Chandrasekhar et al. (2020) who use a model of binary states and repeated guessing. Their experiments use stylized small networks to disentangle the updating rules of subjects. They find that subjects' behaviour is close to that predicted by DeGroot updating.

The empirical literature on networks has highlighted the complex and rich structures and brought out the salience of network homophily and connection inequality. It is unclear what rules of behaviour individuals will follow when confronted by such complex environments. To address this concern, we propose an experiment with large networks that can accommodate key features of empirical networks. This leads us to study three canonical networks: Erdös-Rényi representing a baseline of decentralized contacts (Newman (2018)), Stochastic Block network representing network homophily (see McPherson et al. (2001), Newman (2018)) and Royal Family network capturing highly influential nodes together with local influence (Acemoglu et al. (2011), Bala and Goyal (1998), Mossel et al. (2014)). Our contribution is therefore twofold: one, we propose a new experimental design with canonical networks and two, we show that the learning patterns of our subjects are consistent with predictions of a model where agents follow the DeGroot updating rule.

Our paper is also related to Becker et al. (2017) and Becker et al. (2019) and the ongoing work of Agranov et al. (2020). Specifically, Agranov et al. (2020) consider a star and a core-periphery network, while Becker et al. (2017) study a hub-spoke network. Our paper differs from these papers in the canonical networks we study: these networks are complex and they accommodate salient features of empirical networks like inequality and homophily. To the best of our knowledge, our paper offers the first experimental evidence supporting strong network effects in such a setting and on the consistency of decision making by subjects with DeGroot updating rule.

[1] With observational data, it would be difficult to test these theoretical predictions about network effects because of identification issues. One reason is that network structures are often endogenous and a second reason is that network structures are rarely fully observable in real life; this creates the possibility that there is a gap between what players observe in a network and what a researcher observes. Thus it would be difficult to attribute the change in behaviours to a learning process in a network. Given these concerns with observational data, we resort to controlled laboratory experiments with large-scale networks.

2 Theory and Hypotheses

We use a model with two states, two signals, and two guesses that is taken from Gale and Kariv (2003). There is a set of individuals $N = \{1, 2, ..., n\}$, with $n \geq 2$. There are two possible states of the world, $\omega \in \{0, 1\}$, which individuals believe to be equally likely a priori.

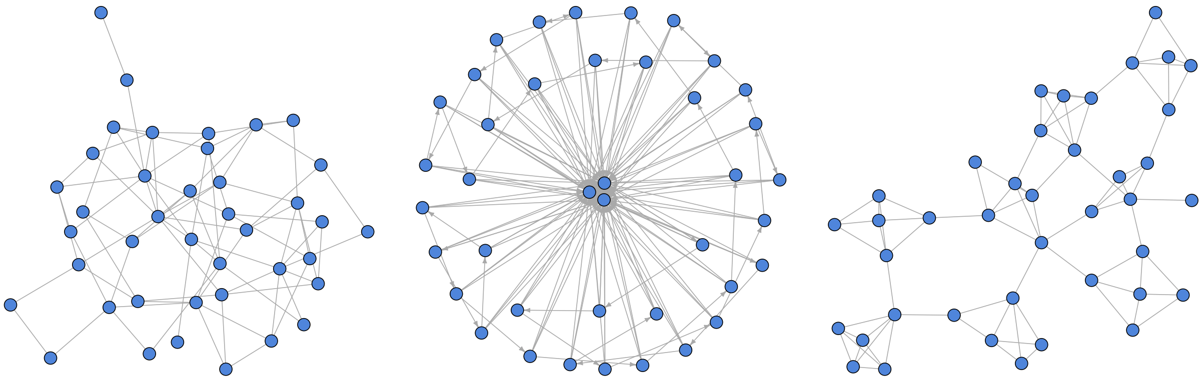
Time is discrete and proceeds as $t = 0, 1, 2....$ In period 0, individuals observe a noisy but informative signal on the true state: individual i receives a binary signal $s_i \in \{0, 1\}$. The probability of receiving the correct signal corresponding to the true state is $p \in (1/2, 1]$. From period $t \geq 1$, an individual chooses a binary guess $a_{i,t} \in \{0, 1\}$. Guessing the true state correctly yields a payoff of 1, and guessing incorrectly yields 0. Thus upon receiving a signal of $s_i = 1$, the expected payoff of an individual guessing $a_{i,t} = 1$ is p and the payoff from guessing $a_{i,t} = 0$ is $1 - p$. Individuals follow their signal in period 1 (note that this guess is also optimal for a myopic individual who seeks to maximise one period payoff).

Individuals are located in an information network, g. We allow for both directed and undirected networks. A link $g_{ij} \in \{0, 1\}$ reflects information access. If $g_{ij} = 1$ then individual i observes the guesses of individual j. $g_{ii} = 0$ by convention. The neighbours of individual i are given by $N_i(g) = \{j | g_{ij} = 1\}$. We will suppose that an individual i gets to observe the guesses of everyone in her neighbourhood. In particular, at time t, individual i observes the guesses of her neighbours from period 1 until period $t-1$. These observations on neighbours' guesses and the signal in period 0 are inputs into individual i's belief at time t about the likelihood of state $\omega = 1$, denoted as $\mu_{i,t}$.

In principle, in period 2, an individual can infer a signal from the first period guess of a neighbour; moreover, in subsequent periods, she can also potentially make inferences on the signals of the neighbours of neighbours, and so forth. These inferences are challenging even in simple situations, but in complex networks, they appear to be even less plausible. With these concerns in mind, building on the literature on majority dynamics (Benjamini et al., 2016) and DeGroot updating (DeGroot, 1974), we propose the following simple rule of thumb for individuals: In period $t = 1$, individual i makes a guess that mimics her signal s_i; in subsequent periods $t \geq 2$, she guesses $a_{i,t}$ that corresponds to the majority guess in her neighbourhood in the previous period (which includes her last period guess $a_{i,t-1}$). To facilitate learning, let us suppose that individuals randomize (with equal probability) between the two states in case of no majority (Grimm and Mengel, 2020). To summarize, an individual i updates her guess $a_{i,t}$ at time t in the following way:

$$a_{i,t} = \begin{cases} 1 & \text{if } \mu_{i,t} > \frac{1}{2}, \\ 0 & \text{if } \mu_{i,t} < \frac{1}{2}, \\ \{0,1\} & \text{if } \mu_{i,t} = \frac{1}{2} \end{cases} \tag{1}$$

$$\text{where} \quad \mu_{i,t} = \frac{1}{|N_i(g)| + 1}\{\sum_{j=1}^{n} a_{j,t-1} \cdot g_{ij} + a_{i,t-1}\}.$$



(a) Erdös-Rényi network (b) Royal Family network (c) Stochastic Block network

Figure 1: Canonical networks

We shall refer to this rule as *DeGroot updating* in the rest of the paper.

We study the learning dynamics and long-run outcomes in three archetypal networks: i) the Erdös-Rényi network; ii) the Stochastic Block network (that reflects network homophily); and iii) the Royal Family network (that represents networks with highly influential individuals and local interaction). Figure 1 presents these networks; we selected these networks as they are representative in their respective classes and have distinct theoretical predictions (see Appendix A.1 for elaborations on the network generation process).

To formulate our hypotheses, we ran simulations of DeGroot learning rule on 1000 sets of signals for each network. The signals are drawn i.i.d. for 40 players with signal quality $p = 0.7$. Players then update their beliefs and guesses under the DeGroot updating rule. We organize the simulation results by defining a variable c_t:

$$c_t = \begin{cases} (n_t - n_0)/(n - n_0) & \text{if } n_t \geq n_0, \\ (n_t - n_0)/n_0 & \text{if } n_t < n_0, \end{cases} \tag{2}$$

where n_0 denotes the number of correct signals received at time 0 and n_t denotes the number of correct guesses made at time t. To account for variations in n_0 (as signals are randomly selected with quality $p = 0.7$), c_t measures the extent to which the average guess at time t move toward correct consensus ($n_t \geq n_0$) or towards incorrect consensus ($n_t < n_0$) relative to the initial assignment of signals. Note that the potential amount of learning towards incorrect consensus is much larger than correct consensus. So the extent of learning is normalized by the maximum margin of learning towards correct consensus ($n - n_0$) or towards incorrect consensus ($n_0 - 0$). Together, c_t ranges between -1 (incorrect consensus) and 1 (correct consensus) with $c_t = 0$ representing no learning.

Figure 2a shows that learning occurs rapidly and the consensus is achieved within the first few periods of the game. This is also reflected in the frequency of switching behaviour: Figure 2b shows that roughly 25% of the individuals switch their guesses in period 2 after observing the guesses of

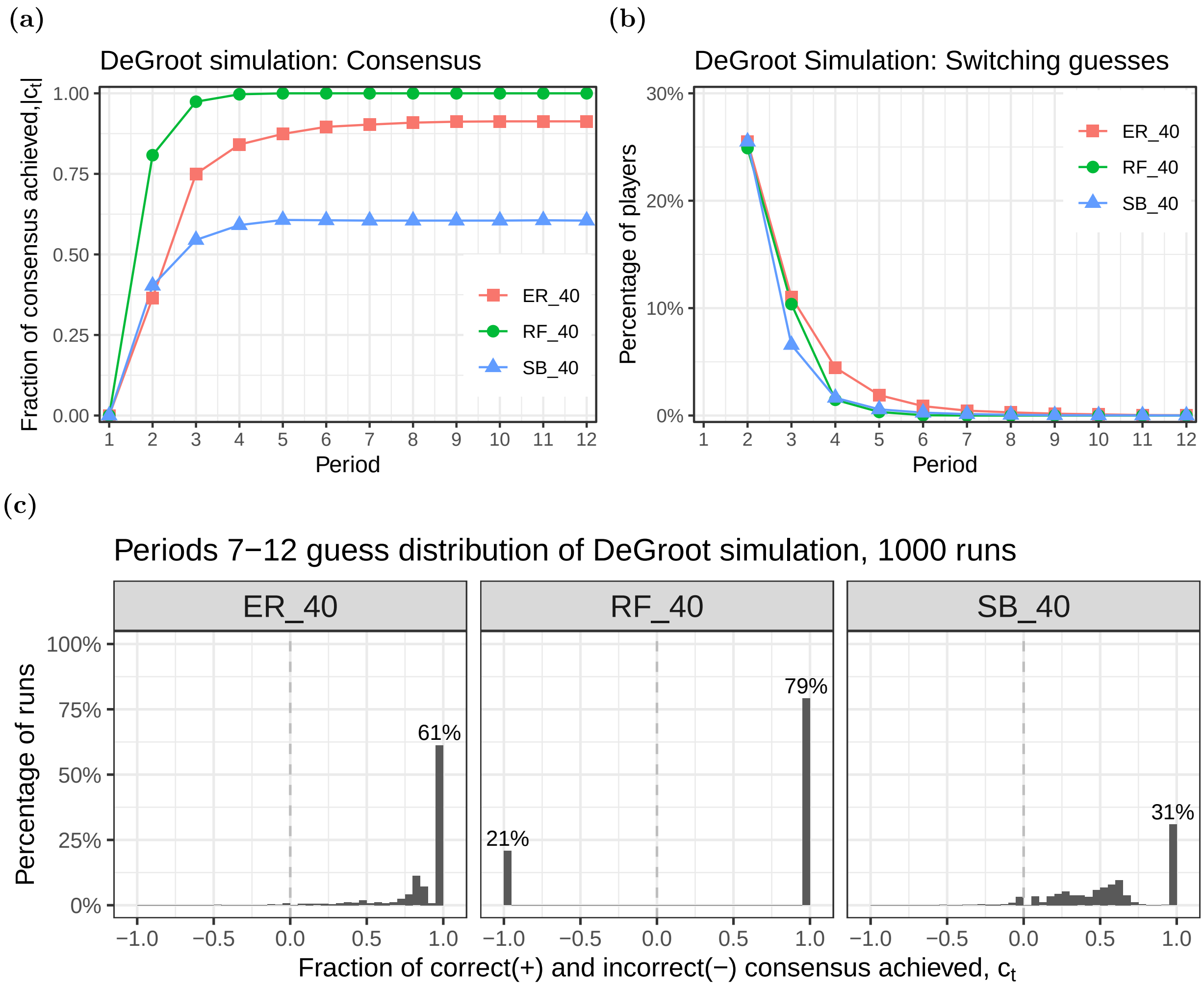


Figure 2: DeGroot simulations. (a) In period 1, $|c_t|$ equals 0 because all individuals guess their signal. By period 4, RF (green) achieves $|c_t| = 100\%$ in almost all cases. SB (blue) attains $|c_t| = 60\%$. ER (red) attains $|c_t| = 87\%$ by period 7. (b) After period 4, less than 5% of individuals switch their guesses from the previous period. By period 7, this frequency is negligible. (c) In periods 7-12, ER network reaches correct consensus in 61% of cases, RF in 79% of cases, and SB in 31% of cases. Almost all remaining cases yield breakdown of consensus in ER and SB (39% and 69%, respectively) or incorrect consensus in RF (21%) (n=1000 per network).

their neighbours. This frequency falls to less than 5% by period 4 and becomes negligible eventually.

We next note that the network has powerful effects on consensus levels. The Royal Family network achieves complete consensus ($c_t = 1$ or -1) by period 4 in almost all simulation runs. By contrast, the Stochastic Block network attains only 60% of potential learning by period 4 and then remains at that level afterwards. Learning in the Erdös-Rényi network continues for longer: the network attains 87% of potential learning by period 7. To separate learning towards correct from learning toward incorrect consensus, Figure 2c presents the distribution of c_t averaged across periods 7-12. In the Erdös-Rényi network, correct consensus obtains in 61% of the cases. In the Royal Family network, consensus obtains in all cases: 79% on correct consensus and 21% on incorrect consensus. In the Stochastic Block model, correct consensus obtains only in 31% of the cases. We obtain similar predictions if we consider variations of the DeGroot updating rule (see Appendix A.3).

We use these theoretical results to formulate three hypotheses:

H.1 Individual guesses converge to a limit guess in all networks.

H.2 The breakdown of consensus is more likely in the Stochastic Block network as compared to the Erdös-Rényi and Royal Family network.

H.3 Incorrect consensus is more likely in the Royal Family network as compared to the Erdös-Rényi and Stochastic Block network.

Let us provide some intuition underlying these hypotheses. The Stochastic Block network is comprised of smaller communities that have a greater density of ties within and fewer ties across them. Since a community is smaller in size than the whole network and has access to fewer signals, it is less likely to reach the correct consensus independently. To illustrate this, consider a scenario where the entire network guesses 1 except for a community that guesses 0. Suppose there is only one link between an individual X (in the community) and the rest of the network, let us say that this link is with individual Y (outside the community). Since X observes herself and other members of her community, she observes a majority guess of 0, while Y observes a majority guess of 1. Under the DeGroot updating rule, X's community therefore agrees upon an incorrect consensus and cannot learn about the external majority (Chandrasekhar et al., 2020). This insulation of communities is more likely in the Stochastic Block than the Erdös-Rényi network because of higher network homophily.

We next discuss why the rate of convergence is higher and why incorrect consensus is so common in the Royal Family network. Observe that, in this network, the 3 members of the 'royal family' (i) constitute a clique among themselves with only one source of information from the outside world, (ii) are observed by everyone in the network, and (iii) constitute a majority in the neighbourhood of everyone. The first property means that the 'royal family' converge to the same guess by period 2. The second and third properties taken together with the DeGroot updating rule imply that everyone outside the 'royal family' imitates the guesses of the 'Royal Family' clique thereby leading to a quick convergence. However, if the majority of the 'Royal Family' happen to get incorrect signals then the consensus will be on the wrong guess.

3 Experimental Design

We recruited 480 participants from the Laboratory for Research in Experimental and Behavioral Economics (LINEEX) at the University of Valencia to take part in a learning game. Subjects were randomized to one of three experimental conditions, each associated with a distinct network structure: Erdös-Rényi, Stochastic Block, and Royal Family network. We ran a total of 12 sessions, 4 sessions for each experimental condition. Each session consisted of a group of 40 subjects on a social network who played 6 rounds of the learning game. No subject participated in more than one session.

In each round of the game, subjects were randomly assigned a position in a social network. Subjects' positions were reshuffled from one round to the next to reduce potential repeated game effects during the experiment (subjects could not keep track of a participant's position across rounds). Subjects in the same session saw the network structure along with different IDs associated with different nodes. Because subjects in the network conditions were not statistically independent, all analyses of collective estimates in the network conditions were conducted at the round level such that each network provided 24 observations. Moreover, because each session completed multiple rounds of the learning game within an experimental trial, we cluster our main analysis at the session level (see Fréchette (2012) for the discussion on dealing with session effects in the laboratory).

Subjects were informed about a bag containing 10 balls. They were told that the bag contains either 7 Red and 3 Green balls (we will refer to this as the RED bag) or 7 Green and 3 Red balls (the GREEN bag). Each of these two combinations is a priori equally likely. At the start of a round, each subject drew a ball from the bag and saw its colour. There was a 70% chance of getting the 'correctly' coloured ball (representing the signal) corresponding to the colour of the bag (representing the true state).

For 12 periods, subjects were asked to guess whether the bag was RED or GREEN. At period $t = 1$, subjects' guess was based on their prior and the colour of the ball initially drawn by them. From period $t \geq 2$ until $t = 12$, subjects also observed guesses of neighbours in previous periods from which they could update their beliefs and revise their guesses. At the end of the round, one period (from 1 to 12) was picked at random to determine actual payoffs in the round: subjects earned 3 euros if their guess matched the colour of the bag (GREEN or RED), and 0 euro otherwise. Total earnings for a subject corresponded to the sum of earnings in each round and a 5 euro show-up fee.

The experiment lasted approximately 1.5 hrs. The average payment per subject was 19.3 euros (including the 5 euro show-up fee). The details of the experimental procedures, including sample instructions, are presented in Appendix D.

4 Findings

We start with a presentation of the learning dynamics. We then compare the level of correct and incorrect consensus and the breakdown of consensus achieved by each network. Lastly, we study whether subjects' behaviour matches various updating rules.

Dynamics of Learning

We begin by discussing the dynamics of learning and the stability of long-run behaviour. Figures 3a and 3b summarize the data. In line with the DeGroot simulation, most of the learning occurs in the early phase of the dynamics: More than three-quarters of the final consensus achieved by period 12 is attained by period 4. In particular, the Royal Family and Stochastic Block networks have more rapid learning than the Erdös-Rényi network. The rapid convergence is also supported by evidence on switching frequency: 20% of subjects switched their guess in period 2 after observing the first-period guess of their neighbours; this switching frequency falls to 10% towards the end of the experiment in period 12. In addition, there are large learning effects across rounds: as a result, the switching probability falls significantly across rounds — only 5% of subjects switched their guess in the last three rounds (Appendix Figure 11). This evidence supports our first hypothesis: *individual guesses converge in all networks.*

Turning to consensus, we note that the level of consensus attained in the experiment is lower than the theoretical prediction (we examine these factors more closely in the Updating Rule section below and in the Appendix). However, the ranking of consensus dynamics across networks is consistent with the DeGroot simulation: the Royal Family network achieves the highest level of consensus from period 2 onward; the Stochastic Block network attains consistently the lowest level of consensus; the Erdös-Rényi network attains level of consensus in between the other two networks.

Consensus Outcomes

We examine the character of long-run outcomes through the measurement of c_t for each network averaged over the last 6 periods, i.e. between periods 7-12, averaged across all rounds and the 4 sessions (similar patterns are obtained if we consider fewer periods or rounds, see Appendix Figure 13). In line with the DeGroot simulation reported in Figure 2c, Figure 3c shows that the distribution in the Royal Family network is bi-modal near $c_t = 1$ and $c_t = -1$, with a higher likelihood on $c_t = 1$ representing correct consensus. The Stochastic Block network has a mode around $c_t = 0$, indicating a greater likelihood of no learning and hence the persistence of diverse opinions. The Erdös-Rényi network leads to a fairly uniform spread of c_t between 0 and 1.

To make a statistical evaluation of the effects of networks on consensus, we proceed as follows: for each round, we average c_t across the last 6 periods. Thus for each network, there are a total of 24 data points (4 sessions with 6 rounds each). Then we categorize each round by whether the averaged

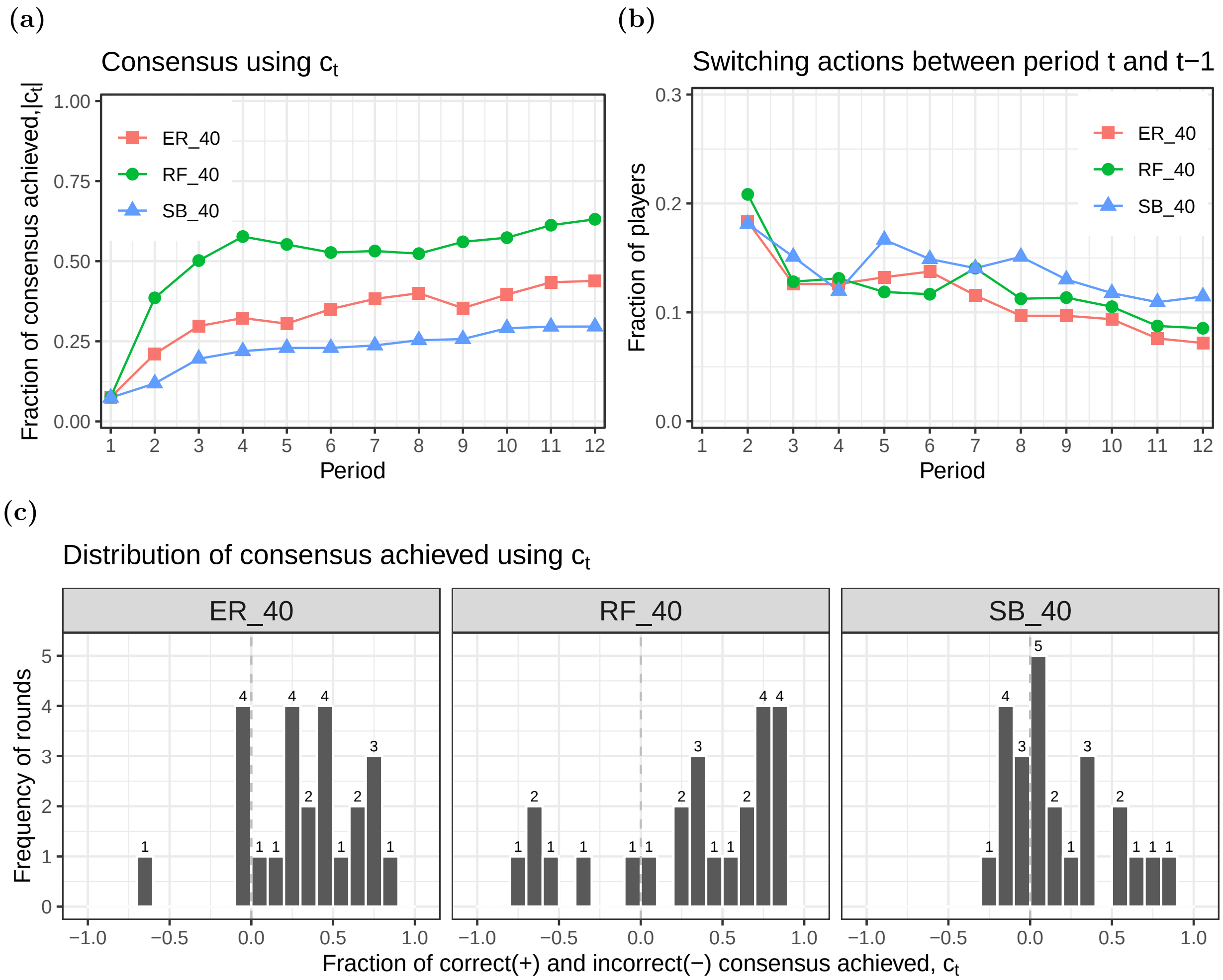


Figure 3: Learning and consensus building. (a) For ER, RF and SB, by period 4, the average $|c_t|$ equals 35%, 58% and 22% respectively. By period 12, ER, RF, and SB, average $|c_t|$ equals 44%, 63%, 30%, respectively. (b) Roughly 20% of subjects switch their guesses in period 2; switching reduces to 10% by period 12. (c) Distribution of c_t is almost uniform between 0 and 1 for ER, bimodal around 1 and -0.7 for RF, and modal around 0 for SB. (n=72: 24 per network).

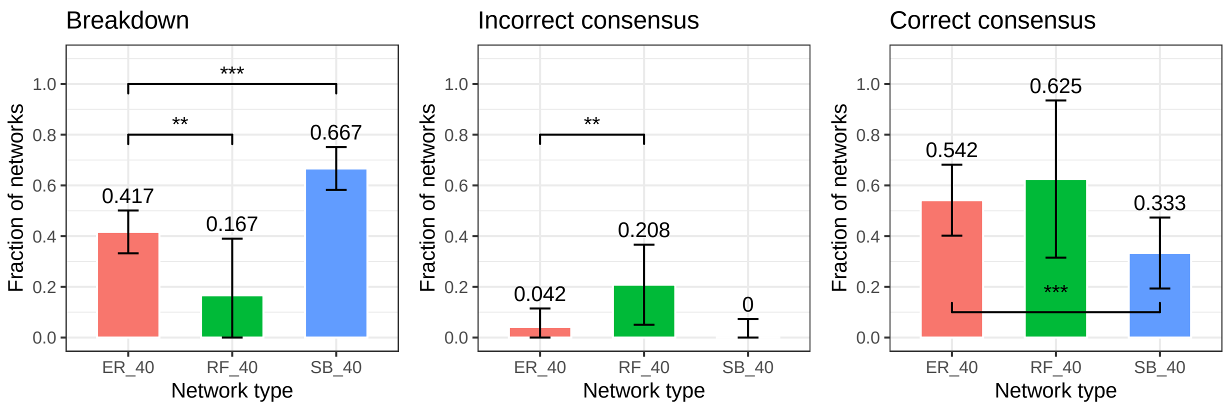


Figure 4: Network effects on consensus. Error bars display standard 95% confidence interval around the mean. Compared to ER: (i) Breakdown of consensus is 25 pp (percentage points) more likely under SB (n=48, 95% CI [0.17,0.33], p-value$<$0.01), and 25 pp less likely under RF (n=48, 95% CI [-0.47,-0.03], p-value$<$0.05); (ii) Incorrect consensus is 17 pp more likely under RF (n=48, 95% CI [0.01,0.32], p-value$<$0.05), and 4 pp less likely under SB (n=48, 95% CI [-0.11,0.03]); (iii) Correct consensus is 21 pp less likely under SB (n=48, 95% CI [-0.35,-0.07], p-value$<$0.01), and 8 pp more likely under RF (n=48, 95% CI [-0.23,0.39]).

c_t is above k (indicating the round achieving correct consensus), below $-k$ (incorrect consensus), or between k and $-k$ (breakdown of consensus). For concreteness, we choose k to be 0.3, so correct consensus is defined as the round achieving more than 30% of the maximum possible learning. Our main findings are robust to different widths k and an alternative, continuous, definition of consensus (Appendix B.2).

In Figure 4, we report the proportion of rounds that achieve correct or incorrect consensus or exhibit a breakdown of consensus for each network (and the corresponding 95% confidence interval). The estimates are derived from the following regression model: for group g in round r,

$$y_{g,r}^{correct} = \beta_0 + \mathbf{1}_g^{RF}\beta_1 + \mathbf{1}_g^{SB}\beta_2 + \epsilon_{g,r}$$

where $\mathbf{1}_g^{RF}$ is an indicator function of whether the group g is playing on the Royal Family network. $y_{g,r}^{correct}$ is an indicator function of whether the round r achieved correct consensus: $\frac{1}{6}\sum_{t=6}^{12} c_{g,r,t} > k$. To account for session effects, we cluster the analysis at the session level (see Fréchette (2012) for the discussion on dealing with session-effects in the laboratory). β_0 can be interpreted as the proportion of Erdös-Rényi networks that reaches correct consensus, whereas β_1 (β_2) can be interpreted as the difference in proportion of networks that reaches correct consensus between Royal Family and Erdös-Rényi network (Stochastic Block and Erdös-Rényi network). Regression results are presented in the Appendix (Table 7).

First, we find that breakdown of consensus is more likely in the Stochastic Block network than the Erdös-Rényi network (n=48, p-value$<$0.01), whereas it is less likely in the Royal Family network (n=48, p-value$<$0.05). Out of 24 rounds and 3 networks (72 data points in total), 22 arrive at breakdown of consensus: 14 in Stochastic Block, 6 in Erdös-Rényi, and 2 in Royal Family network. Recall that there are 8 communities (consisting of 5 individuals each) in the Stochastic Block network. In period 12, 52% of the communities obtain consensus in the Stochastic Block network. This suggests that it is the disagreement across communities that is an important source of the breakdown in consensus in the Stochastic Block network. This is illustrated in 1 round of the Stochastic Block network where more than 7 communities reach complete consensus (5 out of 5 subjects agree) and yet there is breakdown of consensus in the society as a whole. These observations support our second hypothesis: *network homophily leads to breakdown of consensus sustains diverse opinions in a network.*

Second, we find that incorrect consensus is more likely in the Royal Family network than in the Erdös-Rényi network and Stochastic Block network (n=48, p-value<0.05). To appreciate the impact of the 'royal family', note that when 70% of the network receives the correct signal, incorrect consensus is defined as more than half the network guesses incorrectly. Thus the Royal Family network achieves incorrect consensus in 5 rounds (out of 24) as compared to 1 round in Erdös-Rényi and 0 round Stochastic Block network. This supports our third hypothesis: *the presence of highly influential individuals reflected in the Royal Family network, raises the likelihood of incorrect consensus.*

Lastly, we note that correct consensus is less likely in the Stochastic Block network than in Erdös-Rényi (n=48, p-value<0.01) and Royal Family network. Our estimation results are robust to alternative model specifications such as the logit model (Appendix Table 9).

Updating Rule

The environment faced by individuals is complex, so individuals may use different and possibly time-varying updating rules. In this section, we examine how closely individual behaviour matches DeGroot updating.

At every period $t \geq 1$, DeGroot learning predictions are made based on guesses in period $t-1$. We define a binary variable for 'matching DeGroot prediction': it equals 1 when the subject i's guess in period t coincides with the DeGroot prediction, and 0 otherwise. In the case of DeGroot predicting indifference, the variable equals 1 regardless.

Figure 5a presents the percentage of individual guesses that were consistent with DeGroot predictions (in orange colour). We see that, on average, 88% of guesses match with the DeGroot rule (n=11,520 per network). This is higher than the baseline of how well guessing the signal matches with DeGroot predictions: simulations show that only 75% guesses of pseudo subjects (if guessing only their signal) match with DeGroot (see Appendix B.4 for detailed comparison). Figure 5a also presents the fraction of guesses that were contrary to DeGroot prediction but were in line with the signal (in purple colour). In the case when subjects' guesses do not match with DeGroot, about 70% guesses follow their signals. Taken together, DeGroot and persisting with own signal explain more than 95% of the variation in guesses.

By analyzing the fraction of agents that fail to (correctly) guess their signal in period 1, we estimate that about 10% of guesses are made randomly. Indeed, across the networks, the level of consensus achieved in the experiment is comparable to the consensus attained if subjects follow DeGroot updating rule with a 0.1 probability of trembles (see Appendix Figure 17a). This shows that small deviations from DeGroot updating rule at the individual level can have a significant impact on the level of consensus reached. It also suggests that subjects are unlikely to be using other updating rules that are more sophisticated than DeGroot.

Figure 5b presents the time series of the fraction of guesses that matched the prediction of DeGroot updating across rounds. The increase in the match with DeGroot prediction suggests that there is learning across rounds. In particular, as subjects play more rounds, they are more likely to guess their signal in period 1, and they are less likely to persist with their signal in later periods (see Appendix Table 13).

We next turn to heterogeneity in updating rules *across subjects*. The percentage of guesses matching the DeGroot prediction at the subject level is presented in Appendix Figure 18. We see that a substantial fraction of subjects in each network follows DeGroot rule. For instance, 80% of subjects in the Erdös-Rényi network match with DeGroot predictions at least 80% of the time; these fractions are 72% in the Royal Family network and 76% in the Stochastic Block network, respectively. This is again compared to the baseline of how well guessing signal matches with DeGroot predictions: simulations show that only 44% guesses of pseudo subjects (if guessing only signal) in the Erdös-Rényi network match with DeGroot predictions at least 80% of the time (37% in the Royal Family network, and 41% in the Stochastic Block network).

Testing how data matches with other learning rules is generally difficult in large networks. Here we briefly comment on Bayesian learning (for a discussion of variants of DeGroot and other updating rules see Chandrasekhar et al. (2020) and Grimm and Mengel (2020)). As Bayesian rules cannot be computed for large networks, following Chandrasekhar et al. (2020), we consider the role of *information dominant* players. We shall say that player X is an information dominant leader of player Y if X observes Y and all neighbours of Y. A Bayesian player X should ignore guesses of Y (after period 1) while Y should imitate X in all periods. In our experiment, when DeGroot prediction conflicts with the information leader's guess, only around 10% of subjects follow Bayesian prediction (ER:10%, RF:4%, SB:14%), while the rest follow DeGroot prediction. Similarly, when the DeGroot prediction contradicts the signal received, less than 30% of subjects follow their signal (ER:25%, RF:29%, SB:29%), while the rest follow DeGroot. Regression estimates are presented in the Appendix Tables 12 and 13. To sum up, the vast majority of guesses are consistent with the predictions of the DeGroot updating rule.

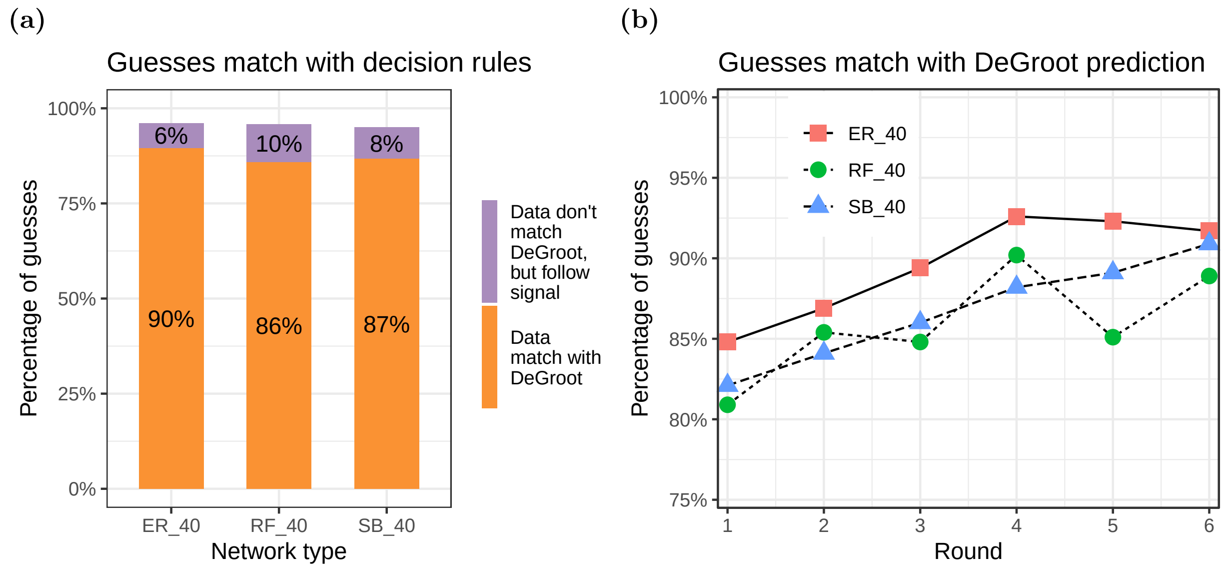


Figure 5: Comparing actual guesses with DeGroot prediction. (a) 88% of guesses match with DeGroot prediction and 6∼10% match with signal. Together they explain 95% of variation in guesses. (b) 80∼85% of guesses match with DeGroot prediction in round 1; this increases to 88∼92% by round 6 (n=34,560: 11,520 per network).

References

Acemoglu, D., M. A. Dahleh, I. Lobel, and A. Ozdaglar (2011). Bayesian learning in social networks. *The Review of Economic Studies 78*(4), 1201–1236. 1

Agranov, M., B. Gillen, and D. Persitz (2020). Information aggregation on networks: an experimental study. *Working paper* .(.), . 1

Bala, V. and S. Goyal (1998). Learning from neighbours. *The review of economic studies 65*(3), 595–621. 1

Bala, V. and S. Goyal (2001). Conformism and diversity under social learning. *Economic theory 17*(1), 101–120. 1

Barabási, A.-L. and R. Albert (1999). Emergence of scaling in random networks. *Science 286*(5439), 509–512. 1

Becker, J., D. Brackbill, and D. Centola (2017). Network dynamics of social influence in the wisdom of crowds. *Proceedings of the national academy of sciences 114*(26), E5070–E5076. 1, A.1

Becker, J., E. Porter, and D. Centola (2019). The wisdom of partisan crowds. *Proceedings of the National Academy of Sciences 116*(22), 10717–10722. 1

Benjamini, I., S.-O. Chan, R. O'Donnell, O. Tamuz, and L.-Y. Tan (2016). Convergence, unanimity and disagreement in majority dynamics on unimodular graphs and random graphs. *Stochastic Processes and their Applications 126*(9), 2719–2733. 2

Breza, E. (2016). Field experiments, social networks and development. In Y. Bramoulle, A. Galeotti, and B. Rogers (Eds.), *Oxford Handbook of the Economics of Networks*. Oxford University Press. 1

Chandrasekhar, A. G., H. Larreguy, and J. P. Xandri (2020). Testing models of social learning on networks: Evidence from two experiments. *Econometrica 88*(1), 1–32. 1, 2, 4, C

Choi, S., D. Gale, and S. Kariv (2005). *Behavioral aspects of learning in social networks: an experimental study*. Emerald Group Publishing Limited. 1

Choi, S., E. Gallo, and S. Kariv (2016). Networks in the laboratory. In Y. Bramoulle, A. Galeotti, and B. Rogers (Eds.), *Oxford Handbook of the Economics of Networks*. Oxford University Press. 1

Coleman, J. S., E. Katz, and H. Menzel (1966). *Medical innovation: A diffusion study*. Indianapolis: Bobbs-Merrill Company. 1

Currarini, S., M. O. Jackson, and P. Pin (2009). An economic model of friendship: Homophily, minorities, and segregation. *Econometrica 77*(4), 1003–1045. 1

DeGroot, M. H. (1974). Reaching a consensus. *Journal of the American Statistical Association 69*(345), 118–121. 1, 2

DeMarzo, P. M., D. Vayanos, and J. Zwiebel (2003). Persuasion bias, social influence, and unidimensional opinions. *The Quarterly journal of economics 118*(3), 909–968. 1

Fréchette, G. R. (2012). Session-effects in the laboratory. *Experimental Economics 15*(3), 485–498. 3, 4

Gale, D. and S. Kariv (2003). Bayesian learning in social networks. *Games and economic behavior 45*(2), 329–346. 1, 2

Galton, F. (1907). Vox populi (the wisdom of crowds). *Nature 75*(7), 450–451. 1

Golub, B. and M. O. Jackson (2010). Naive learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics 2*(1), 112–149. 1

Golub, B. and E. Sadler (2016). Learning in social networks. In Y. Bramoulle, A. Galeotti, and B. Rogers (Eds.), *Oxford Handbook of the Economics of Networks*. Oxford University Press. 1

Goyal, S. (forthcoming). *Networks: An Economics Approach.* MIT Press. 1

Grimm, V. and F. Mengel (2020). Experiments on belief formation in networks. *Journal of the European Economic Association 18*(1), 49–82. 1, 2, 4

Katz, E. and P. F. Lazarsfeld (1966). *Personal Influence, The part played by people in the flow of mass communications.* Transaction Publishers. 1

Kearns, M., S. Judd, and Y. Vorobeychik (2012). Behavioral experiments on a network formation game. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pp. 690–704. 1

Lazarsfeld, P. F. and R. K. Merton (1954). Friendship as a social process: A substantive and methodological analysis. *Freedom and control in modern society 18*(1), 18–66. 1

McPherson, M., L. Smith-Lovin, and J. M. Cook (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology 27*(1), 415–444. 1

Mobius, M., T. Phan, and A. Szeidl (2015). Treasure hunt: Social learning in the field. Technical report, National Bureau of Economic Research. 1

Mossel, E., A. Sly, and O. Tamuz (2014). Asymptotic learning on bayesian social networks. *Probability Theory and Related Fields 158*(1-2), 127–157. 1

Newman, M. (2010). *Networks: An Introduction.* New York, NY, USA: Oxford University Press, Inc. 1

Newman, M. (2018). *Networks.* Oxford university press. 1

Appendix: For Online Publication

A Simulation

A.1 Network generation & selection

There are 3 networks of interest: Erdös-Rényi (ER), Stochastic Block (SB), and Royal Family (RF) network. All networks have 40 nodes, $n = 40$. In order to control for the average information received of each node, the networks have an average outdegree of 4 (excluding self links), following Becker et al. (2017). Our DeGroot simulations show that the hypotheses are robust against different outdegrees.

The generation process of each network type is as follows. The parameter specifications in the network generation process were selected to ensure strong connectedness in the networks generated.

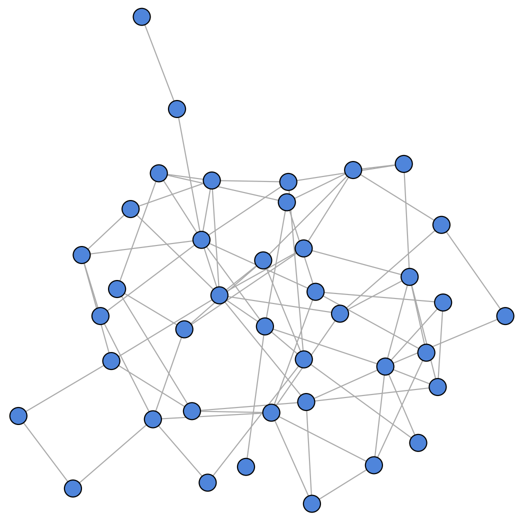
- Erdös-Rényi networks are generated according to the Erdös-Rényi model (using the "erdos.renyi.game" function from the igraph package). We specify the number of nodes as n and total number of edges as $2n$.
- Stochastic Block networks are generated according to the Trait-based random generation (using "sample_pref" function from the igraph package). We specify the number of nodes as n and the size of each community as 5. So there are $n/5$ communities where the probability of linking within a community is $p_{ii} = 0.85$ and between communities is $p_{ij} = p_{ii}/60$. (These parameter specifications were selected to ensure strong connectedness in the networks generated)
- Royal Family networks are created by first placing n players in a directed ring (player n observes player 1 who then observes player 2 and so on). Then players 1,2,3 are selected to be the hub where all players observe them. All players have outdegree of 4 (except for player 1 and 2 with outdegree of 2, and player 3 and n with outdegree of 3).

For each network treatment, we randomly generated 100 networks that are (strongly) connected — every node can be reached through a path from every other node. Then we computed network measures such as outdegree, diameter, average path length, clustering for each network. The average statistics for each network type are presented in Table 1.

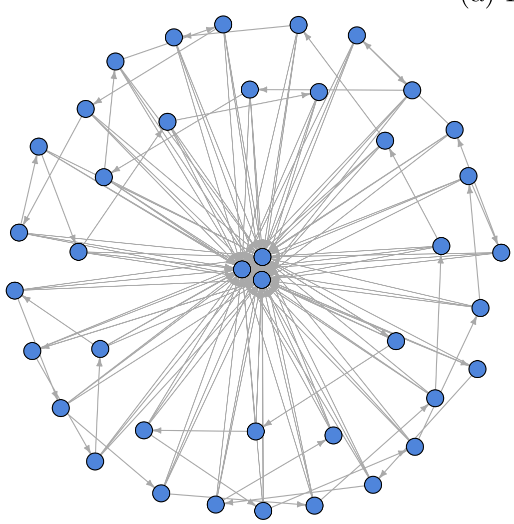
Out of the 100 randomly generated networks, a network with measures closest to the average statistics is then used in the experiment. Table 2 presents the network statistics of these networks and Figure 6 presents the network graphs. Note that the Royal Family network is not generated randomly.

A.2 Signal generation & selection

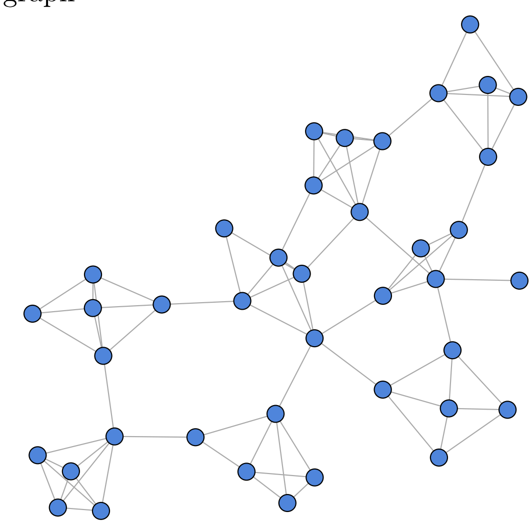
We randomly select 24 sets of signals for the experiment. For each network treatment, there are 4 groups of players each playing 6 rounds. So group 1 in round 1 uses the first set of signals while



(a) ER graph



(b) RF graph



(c) SB graph

Figure 6: Network graph of n=40 with average outdegree 4

Table 1: Averages network statistics of 100 randomly generated networks

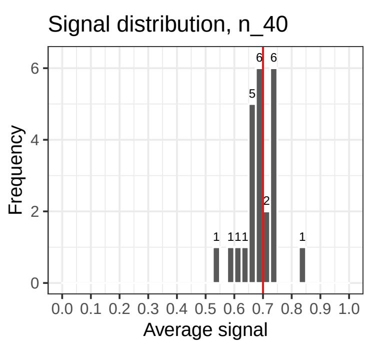
n=40	avg. outdegree	diameter	avg path length	clustering
ER	4.00	5.63	2.73	0.10
SB	3.98	9.15	4.12	0.57
RF	3.85	38.00	12.72	0.26

Table 2: Network statistics of the networks used in the experiments

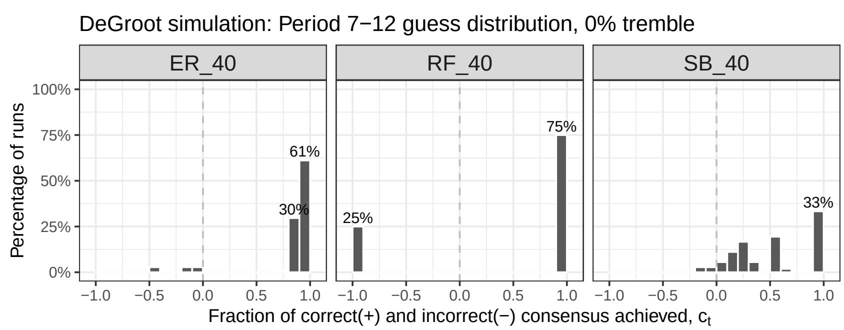
n=40	avg. outdegree	diameter	avg path length	clustering
ER	4.00	5	2.73	0.10
SB	4.00	9	3.85	0.57
RF	3.85	38	12.72	0.26

group 4 in round 6 uses the 24th set. Therefore, the same collection of signals are used across all networks.

We perform two checks to ensure that the 24 sets of signals are representative. First, we note that the distribution of the 24 sets of signals is bell-shaped around the mean 0.7 where 1 represents the correct state (Figure 7a). Second, we confirm that the simulated guesses following these 24 sets of signals (Figure 7b) have the same properties as the simulations of the 1000 sets of signals (presented in Figure 2c, see main text). The regression on network effects with respect to 'Correct consensus', 'Incorrect consensus' and 'Breakdown of consensus' (as defined in the Consensus Outcomes section in the main text) confirms the main hypotheses (Table 3).



(a) Distribution of signals



(b) Distribution of c_t simulated from signals

Figure 7: Signal distribution and simulation results using the 24 sets of signals for the experiment. (a) Distribution of signals used for all networks in the experiment with mean 0.70, standard deviation 0.06, 1st quartile 0.675, 2nd quartile 0.70 and 3rd quartile 0.75. (n=24) (b) Distribution of c_t under DeGroot simulation using experiment signals. The hypotheses from the simulation of 1000 runs are confirmed: 1) There is more breakdown of consensus in the Stochastic Block network than in the Erdös-Rényi and Royal Family network; 2) There is more incorrect consensus in the Royal Family network than in the Erdös-Rényi and Stochastic Block network.

Table 3: OLS regression of simulated data, network size 40, k = 0.3

	OLS - Correct Consensus	OLS - Incorrect Consensus	OLS - Breakdown
(Intercept)	0.88^{***}	0.04	0.08
	(0.09)	(0.05)	(0.07)
typeRF	-0.08	0.17^{**}	-0.08
	(0.12)	(0.08)	(0.10)
typeSB	-0.37^{***}	-0.04	0.42^{***}
	(0.12)	(0.08)	(0.10)
R^2	0.13	0.11	0.31
Adj. R^2	0.10	0.08	0.29
Num. obs.	72	72	72

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

A.3 Variations on DeGroot updating rule

All network effects identified in the simulations are robust to alternative variations on DeGroot updating rule.

Deterministic DeGroot. In the case of indifference, suppose an individual persists with her last period's guess. Formally, we say:

$$a_{i,t} = \begin{cases} 1 & \text{if } \mu_{i,t} > \frac{1}{2}, \\ 0 & \text{if } \mu_{i,t} < \frac{1}{2}, \\ a_{i,t-1} & \text{if } \mu_{i,t} = \frac{1}{2} \end{cases} \tag{3}$$

Simulations of this variant of the DeGroot are presented in Figure 8.

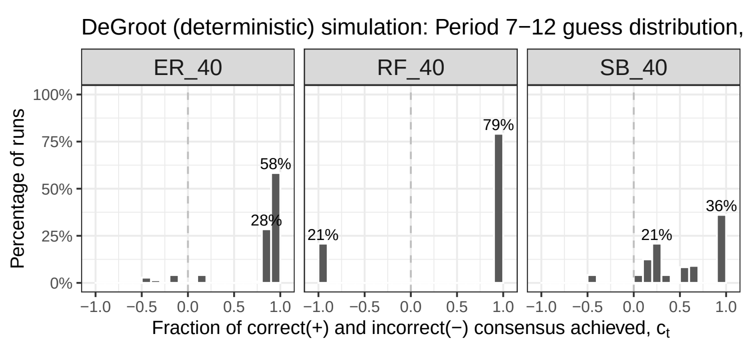


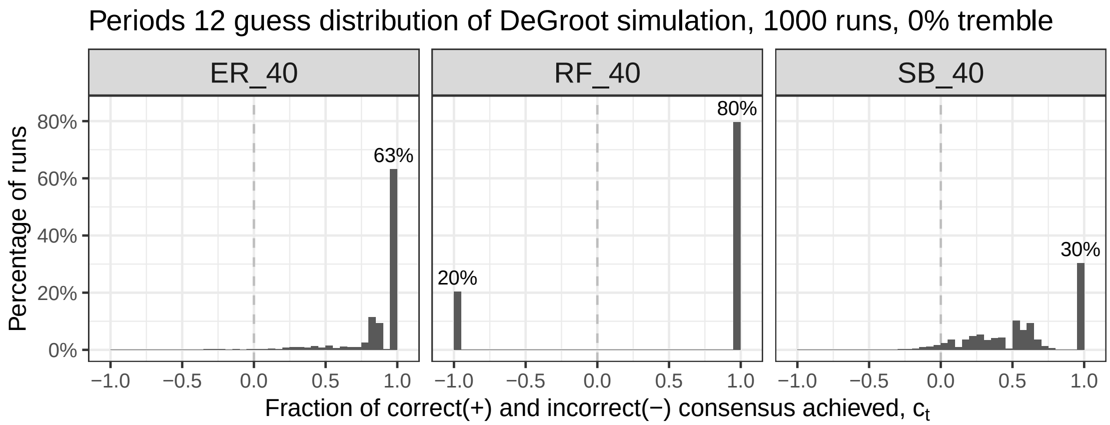
Figure 8: Distribution of c_t under Deterministic DeGroot simulation using experiment signals.

DeGroot with Trembling. Suppose an individual observes a majority guess of Red: if we use DeGroot updating rule with 10% trembling, that means she would guess Green 10% of the time and

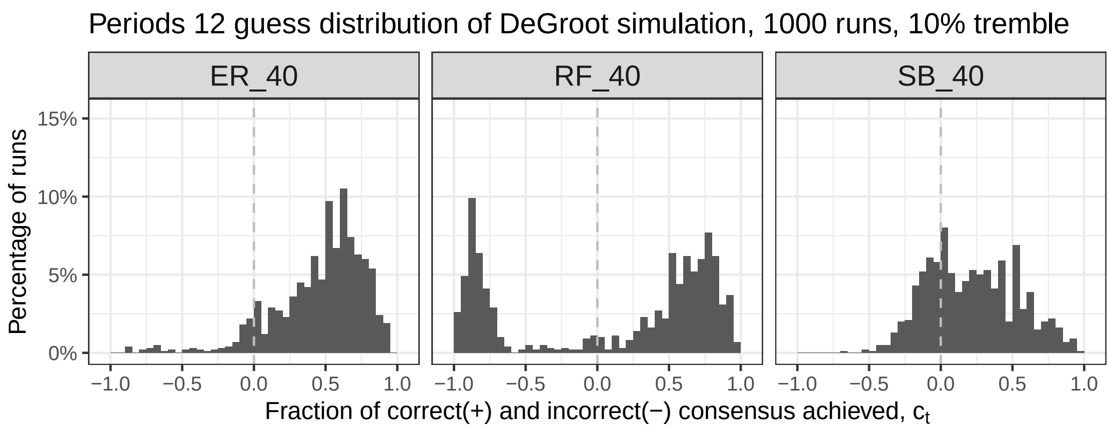
Red 90% of the time. Figure 9 shows that the networks effects identified with the original DeGroot (as in Figure 2c in the main text) are robust.

A.4 Welfare

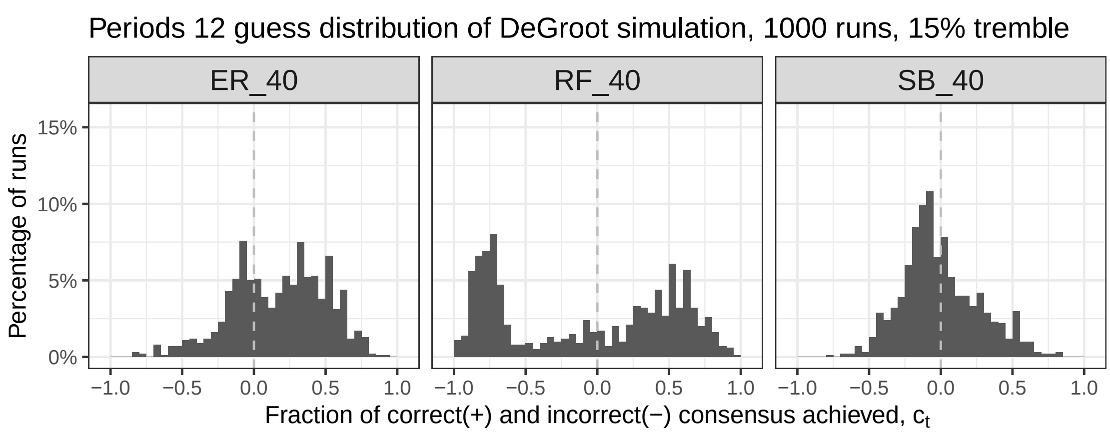
To compare welfare performances across networks, we use a simple measure of welfare improvement: $w_t = (n_t - n_0)/(n - n_0)$, i.e., the first part of c_t in eq. (2), applied to the entire range. The measure ranges from 1 (maximum gain in welfare) to $-n_0/(n-n_0)$ (maximum loss in welfare). The simulation shows that, in the limit, Erdös-Rényi has the highest average welfare improvement, followed by Stochastic Block, and Royal Family comes last (Figure 10). The reason for the poor showing of the Royal Family network is the high frequency of incorrect consensus outcomes. Stochastic Block network achieves lower welfare improvement than Erdös-Rényi network because it has more breakdowns of correct consensus.



(a) 0% trembling



(b) 10% trembling



(c) 15% trembling

Figure 9: Distribution of c_t under simulation with trembling.

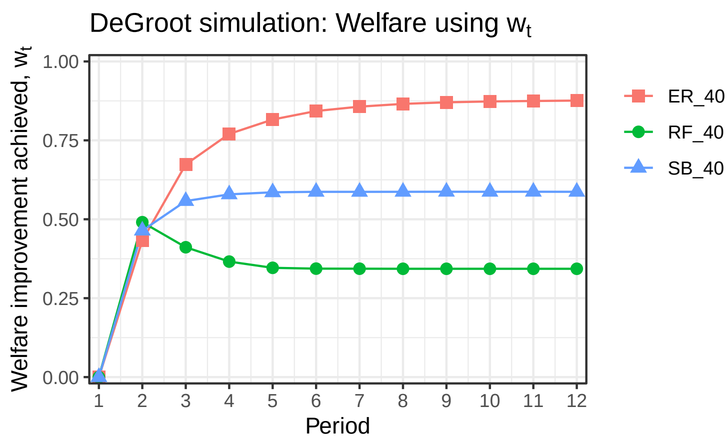
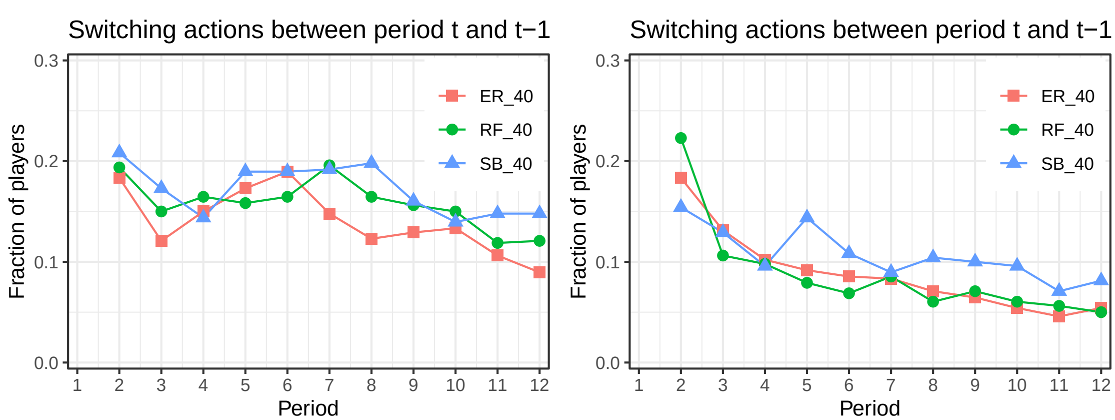


Figure 10: Evolution of welfare improvement w_t in the simulation: In period 12, ER achieved 88% of the possible welfare improvement, 58% for SB and 34% for RF.

B Findings

B.1 Convergence

The rapid convergence of guesses in the experiment is supported by evidence on switching frequency: 20% of individuals switched their guesses at period 2 after observing the first period guesses of their neighbors, this switching frequency falls to 10% toward the end of the experiment in period 12. The switching probability falls significantly as subjects learn across rounds: as a result, it is only 5% in the last three rounds (Figure 11). We argue that the residual switching in guesses in the final periods are not due to further learning by subjects, but due to random guessing.



(a) n=40, round 1-3 (b) n=40, round 4-6

Figure 11: Percentage of subjects switching guesses per period. (a) In the first three rounds, percentage of switching falls from 20% in period 2 to 10∼15% in period 12. (b) In the last three rounds, percentage of switching falls from around 20% in period 2 to 5∼8% in period 12. Therefore, adjusting for learning across rounds, there is less than 8% of subjects switching guesses by period 12.

We estimate that 10% of the guesses are random in the experiment, using the following technique: Irrespective of whether a myopic player follows Bayesian or DeGroot learning rule, in period 1, it is optimal to guess her initial signal. In period 2, both (myopic) Bayesian and DeGroot learning rules predict that player should follow the majority guess in her neighbourhood in period 1. Table 4 shows that about 10% of guesses do not follow subjects' initial signals in period 1 and contradict both learning rules in period 2. This suggests that about 10% of guesses ignore information.

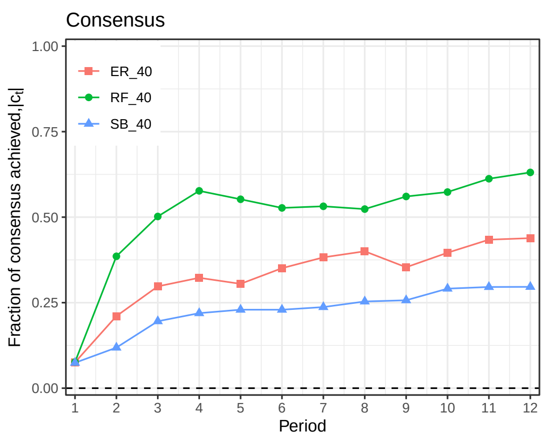
Table 4: Fraction of guesses against Bayesian and DeGroot prediction, network size 40

	Guess against majority in period 1,2	
	OLS (Bayesian, DeGroot predicts 0)	Logit
(Intercept)	0.10^{***}	-2.24^{***}
	(0.01)	(0.08)
typesizeRF_40	0.02^{*}	0.24^{*}
	(0.01)	(0.13)
typesizeSB_40	0.02^{***}	0.25^{***}
	(0.01)	(0.08)
R^2	0.00	
Adj. R^2	0.00	
Num. obs.	5760	5760
AIC		4029.57
BIC		4049.54
Log Likelihood		-2011.78
Deviance		4023.57

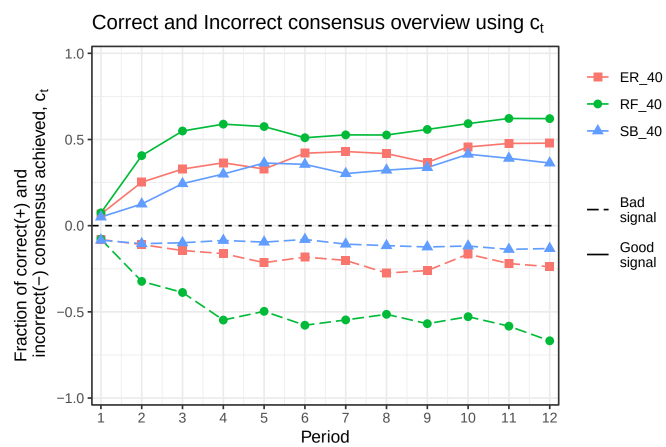
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

B.2 Consensus

The simulations lead us to propose two hypotheses: One, the breakdown of consensus is most likely in the Stochastic Block network, followed by the Erdös-Rényi network and lastly the Royal Family network; Two, the Royal Family network leads to the wrong consensus more often than the Erdös-Rényi network. Figure 12a presents the evolution of consensus across periods across all networks, while Figure 12b presents the evolution of c_t partitioned by 'good' and 'bad' signals. Under DeGroot updating simulation, the set of 'good' signals would lead to $c_t \geq 0$ (correct consensus), while the 'bad' signals would lead to $c_t < 0$ (incorrect consensus). They show that the rankings in the hypotheses are maintained across all periods. The regression Table 7 shows the statistical significance of the estimates (presented in Figure 4 in the main text), supporting our hypotheses. The estimate of 'incorrect consensus' on 'typeRF' represents the difference in fraction of incorrect consensus achieved between the Royal Family network and the Erdös-Rényi network.



(a) Evolution of consensus $|c_t|$



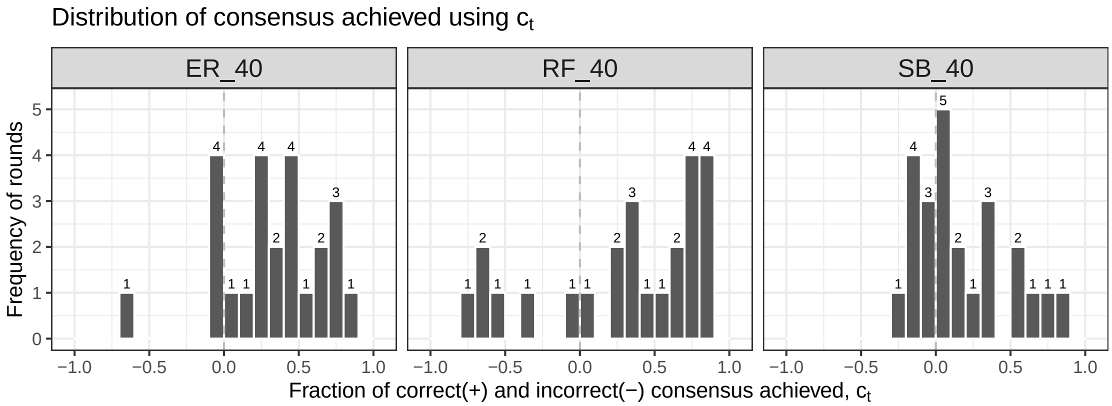
(b) Evolution of consensus c_t

Figure 12: Evolution of $|c_t|$ and partitioned c_t. (a) In period 12, RF, ER, SB reach 63%, 44%, 30% of consensus, respectively. (b) We partitioned c_t averaged across all games by 'good' and 'bad' signals. The ranking of correct and incorrect consensus reached is preserved across most periods.

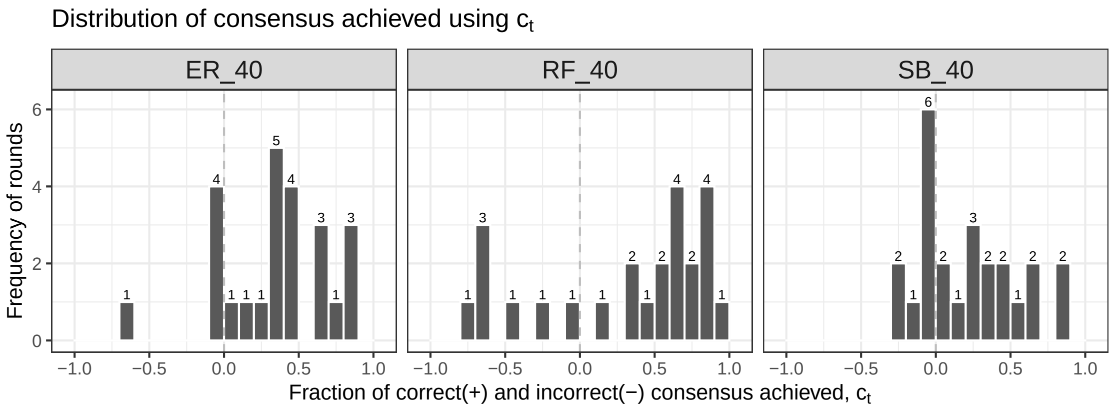
A similar distribution of c_t obtains if we consider fewer periods (periods 10-12) or rounds (rounds 4-6) (Figure 13).

Recall, we defined binary variables of correct consensus (if $c_t > k$), incorrect consensus (if $c_t < -k$), and breakdown of consensus (if $-k \leq c_t \leq k$) based on the value of c_t. Our main findings are robust to 1) different widths k (Tables 6 to 8), 2) an alternative model specification such as the logit model (Table 9), and 3) a continuous definition of consensus outcomes (Table 10).

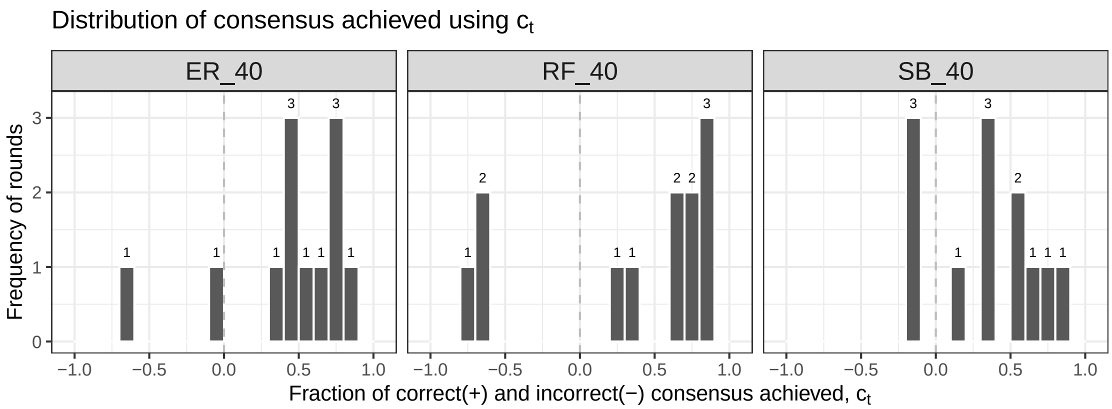
A continuous variation on the definition of consensus would be as follows: Consensus is defined as the absolute value of c_t, $|c_t|$; correct consensus is defined as censoring negative values of c_t to 0; incorrect consensus censors positive values of c_t to 0; breakdown is defined as the negative of consensus, $-|c_t|$.



(a) Distribution of averaged c_t, between period 7-12, round 1-6 (n=24 per network)



(b) Distribution of averaged c_t, between period 10-12, round 1-6 (n=24 per network)



(c) Distribution of averaged c_t, between period 7-12, round 4-6 (n=12 per network)

Figure 13: Distribution of averaged c_t robust over period and round selections.

For the Stochastic Block network, we show that communities are more likely to reach consensus compared to Erdös-Rényi network despite the networks being less likely to reach consensus. The subset of subjects in each block of the Stochastic Block model may be seen as constituting a 'community'. Given the network generation methods in the Stochastic Block model, subjects with location id $1-5$ is a community while id $6-10$ is another community, and so on. So in all three networks, we define a community by the same location ids. We define *community consensus* as 1 when all 5 subjects in a community reaches complete consensus, and 0 otherwise.

Table 5 shows that 52% of communities reach consensus in the Stochastic Block network which is 14% point higher than in Erdös-Rényi network. We next look closer at the dispersion of average guesses of communities. The maximum difference in average guesses of communities is equal to 1: when there exists one community with correct consensus and one with incorrect consensus. Figure 14 shows that 75% of rounds in the Stochastic Block network have large dispersion in community guesses (greater than 0.7) while only 50% in Erdös-Rényi network and 46% in Royal Family network. This implies that disagreements between communities are the principal source of the consensus breakdown in the Stochastic Block network.

B.3 Welfare

We examine the implications of learning outcomes on welfare by using the welfare improvement measure w_t. In line with the DeGroot simulation, the Erdös-Rényi network has higher welfare improvement than the Stochastic Block network (Table 11). We find that the Erdös-Rényi network achieves the average welfare improvement of 0.30, while only 0.18 for the Stochastic Block network. We do not find any significant difference in welfare improvement between the Erdös-Rényi network and the Royal Family network. This is probably due to the large variation in welfare improvement and deterioration among the Royal Family networks.

Table 5: Regression of community consensus on network treatment

	OLS - Community Consensus	Logit - Community Consensus
(Intercept)	0.38***	−0.49***
	(0.03)	(0.15)
typeRF	0.16*	0.66*
	(0.09)	(0.37)
typeSB	0.14***	0.55***
	(0.04)	(0.16)
R^2	0.02	
Adj. R^2	0.02	
Num. obs.	576	576
AIC		791.85
BIC		804.92
Log Likelihood		−392.93
Deviance		785.85

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

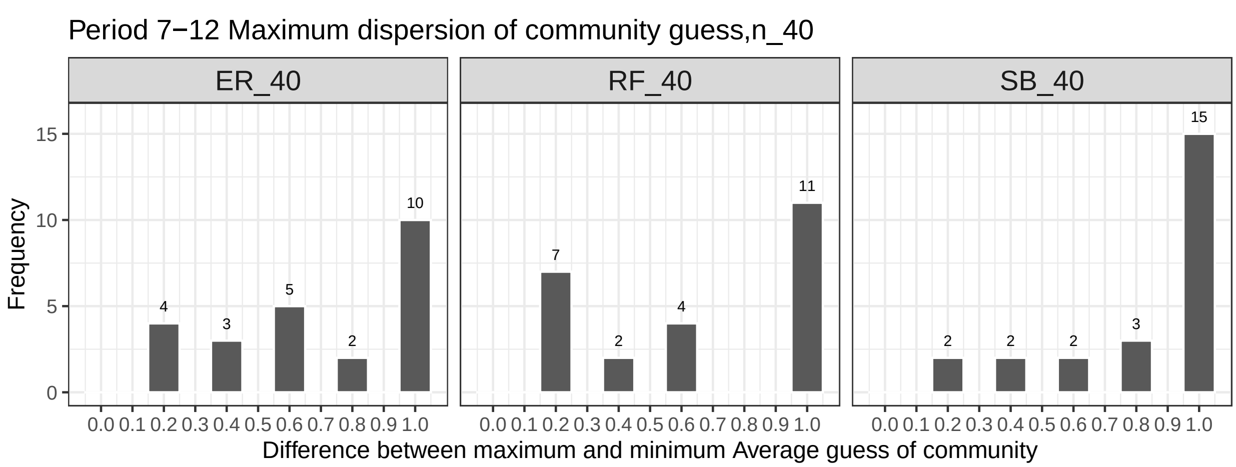


Figure 14: Distribution of the maximum dispersion in average guesses between communities for each network. 75% of rounds in the SB have more than 0.7 dispersion in average guesses between communities, 50% in ER and 46% in RF (n=24 per network).

Table 6: OLS regression c_t, k=0.2, n=40

	OLS - Correct Consensus	OLS - Incorrect Consensus	OLS - Breakdown
(Intercept)	0.71^{***}	0.04	0.25^{***}
	(0.07)	(0.04)	(0.04)
typeRF	-0.00	0.17^{**}	-0.17^{*}
	(0.15)	(0.08)	(0.09)
typeSB	-0.33^{***}	-0.00	0.33^{***}
	(0.08)	(0.05)	(0.06)
R^2	0.10	0.07	0.20
Adj. R^2	0.08	0.04	0.18
Num. obs.	72	72	72

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 7: OLS regression c_t, k=0.3, n=40

	OLS - Correct Consensus	OLS - Incorrect Consensus	OLS - Breakdown
(Intercept)	0.54^{***}	0.04	0.42^{***}
	(0.07)	(0.04)	(0.04)
typeRF	0.08	0.17^{**}	-0.25^{**}
	(0.16)	(0.08)	(0.11)
typeSB	-0.21^{***}	-0.04	0.25^{***}
	(0.07)	(0.04)	(0.04)
R^2	0.06	0.11	0.17
Adj. R^2	0.03	0.08	0.15
Num. obs.	72	72	72

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 8: OLS regression c_t, k=0.4, n=40

	OLS - Correct Consensus	OLS - Incorrect Consensus	OLS - Breakdown
(Intercept)	0.46^{***}	0.04	0.50^{***}
	(0.09)	(0.04)	(0.06)
typeRF	0.04	0.12^{*}	-0.17
	(0.18)	(0.07)	(0.12)
typeSB	-0.25^{**}	-0.04	0.29^{***}
	(0.12)	(0.04)	(0.09)
R^2	0.07	0.08	0.14
Adj. R^2	0.04	0.05	0.12
Num. obs.	72	72	72

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 9: Logistic regression c_t, k=0.3, n=40

	Logit - Correct Consensus	Logit - Incorrect Consensus	Logit - Breakdown
(Intercept)	0.17	-3.14^{***}	-0.34^{*}
	(0.28)	(0.92)	(0.17)
typeRF	0.34	1.80^{*}	-1.27^{*}
	(0.66)	(1.01)	(0.77)
typeSB	-0.86^{***}	-16.43^{***}	1.03^{***}
	(0.28)	(1.05)	(0.17)
AIC	101.41	38.88	90.78
BIC	108.24	45.71	97.61
Log Likelihood	-47.71	-16.44	-42.39
Deviance	95.41	32.88	84.78
Num. obs.	72	72	72

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 10: OLS regression c_t censored, n=40

	OLS - Correct Consensus	OLS - Incorrect Consensus	OLS - Breakdown
(Intercept)	0.36^{***}	0.02	0.62^{***}
	(0.03)	(0.02)	(0.01)
typeRF	0.07	0.08^{**}	-0.15^{*}
	(0.10)	(0.04)	(0.08)
typeSB	-0.14^{***}	-0.02	0.16^{***}
	(0.04)	(0.02)	(0.02)
R^2	0.08	0.09	0.19
Adj. R^2	0.06	0.06	0.17
Num. obs.	72	72	72

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 11: OLS regression of welfare w_t on network treatment

	OLS - Welfare
(Intercept)	0.30***
	(0.07)
typeRF	−0.12
	(0.17)
typeSB	−0.18**
	(0.08)
R^2	0.02
Adj. R^2	0.01
Num. obs.	432

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

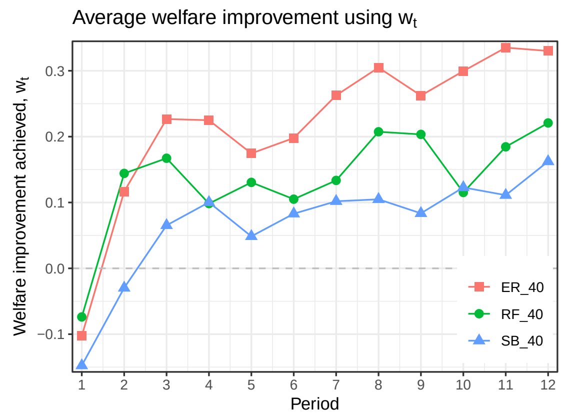


Figure 15: Evolution of welfare improvement w_t in the experiment. In the experiment, in the last period, ER achieved 30% of the possible welfare improvement, 18% for SB and 22% for RF.

B.4 Updating rule

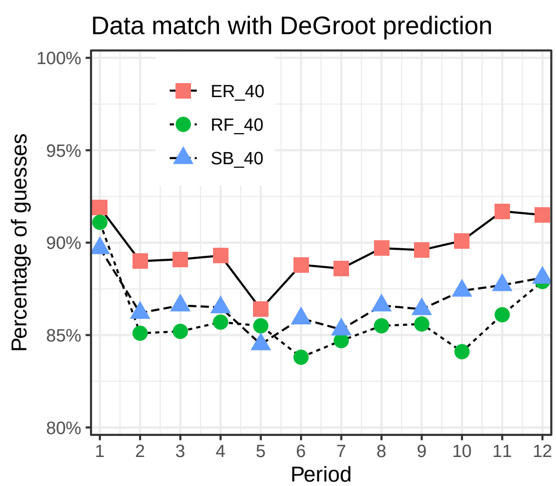


Figure 16: The percentage of guesses matching DeGroot prediction across periods. Across all networks at least 90% of first guesses matched with the DeGroot prediction (i.e., guess follows the signal). This percentage falls to 80%∼85% in the second period and then steadily increases until it reaches 85%∼90% in later periods.

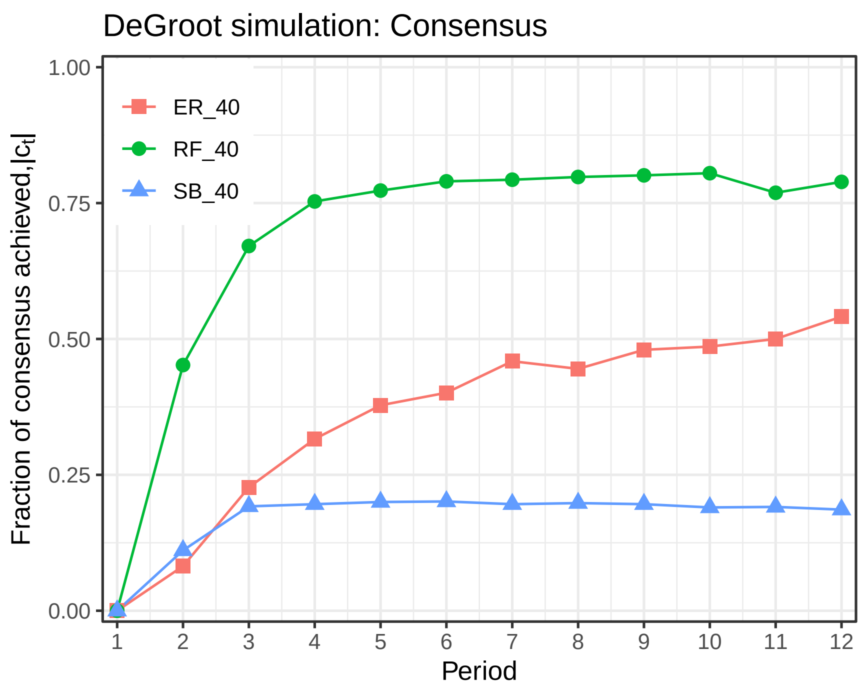
On average, 88% of guesses match with the DeGroot rule. This is higher than the baseline of how well *guessing randomly* matches with DeGroot predictions: simulations show that on average 60% pseudo subjects' random guesses match with DeGroot. This is also higher than the baseline of how well *guessing signal* matches with DeGroot predictions: simulations show that on average 75% guesses of pseudo subjects (if guessing only signal) match with DeGroot (Figure 18a).

Suppose that 10% of guesses are randomly made. We show that the level of consensus attained in the experiment is comparable the simulation under 10% trembling for Erdös-Rényi (Figure 17a) and 15% trembling for Stochastic Block and Royal Family network (Figure 17b).

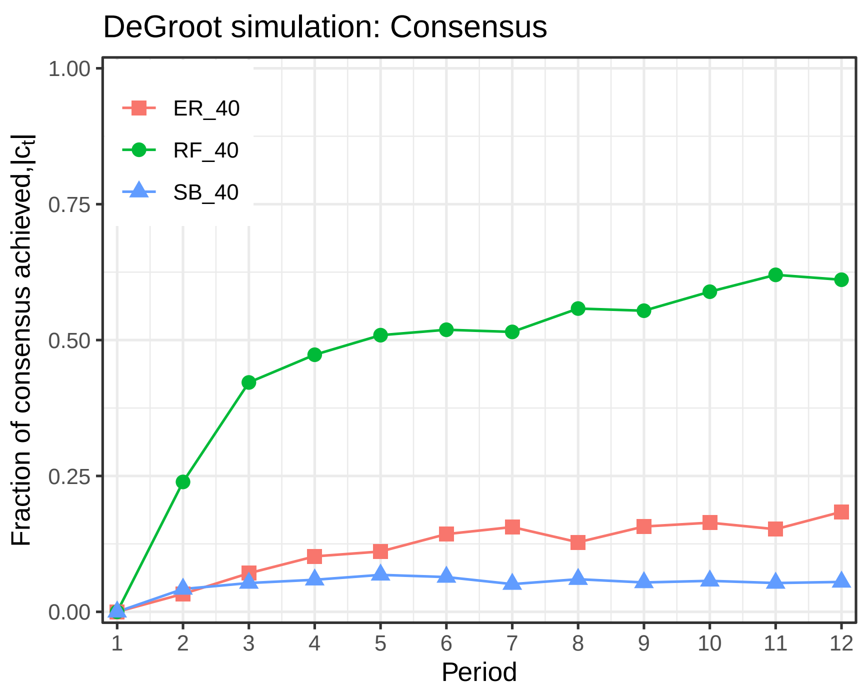
We next delve deeper by looking at subject level match with DeGroot. Because each subject plays a total of 6 rounds and 12 periods per round and their guesses are not statistically independent, we treat each subject as a data point. Figure 18 presents the histogram of how well a subject's guesses match with DeGroot predictions. For all networks, there are significantly more subjects whose guesses match with DeGroot than pseudo subjects who guess their signals or randomly.

Bayesian learning: Information Leader. When DeGroot prediction contradicts with information leader's guess, a Bayesian player should follow their information leader while a DeGroot player should follow the majority of their neighbours. Table 12 show that when the two are in conflict, only around 10% of subjects follow Bayesian prediction (ER:10%, RF:4%, SB:14%), while the rest follow DeGroot. Note that this percentage is decreasing in rounds which suggests subjects' learning behaviour towards DeGroot updating rule.

No learning: Stubborn players that only follow their signal. Similarly, when DeGroot prediction contradicts goes against initial signal received, a stubborn player should only follow their own signal. Table 13 show that around 25% of subjects follow initial signal (ER:25%, RF:29%, SB:29%) while the rest follow DeGroot. As before, we show that there is significant learning across rounds. More interestingly, increases in periods also decreases stubbornness. This could be due to increasing availability of information and therefore less weight is put on initial signals.

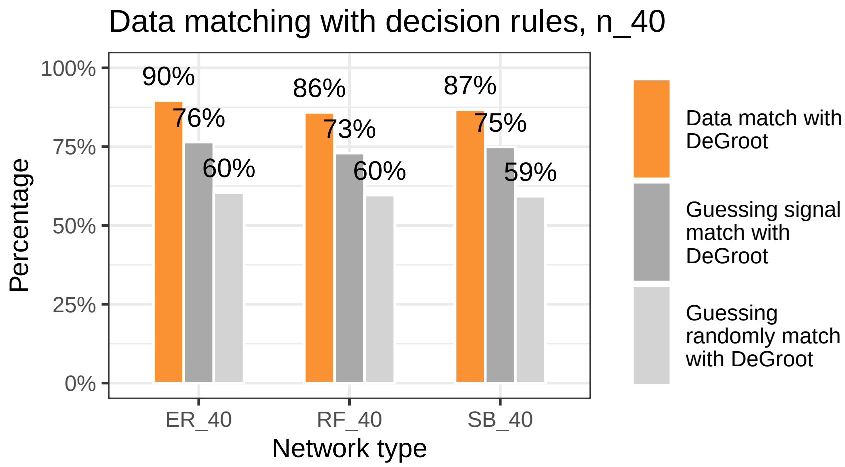


(a) DeGroot simulation with 10% trembling

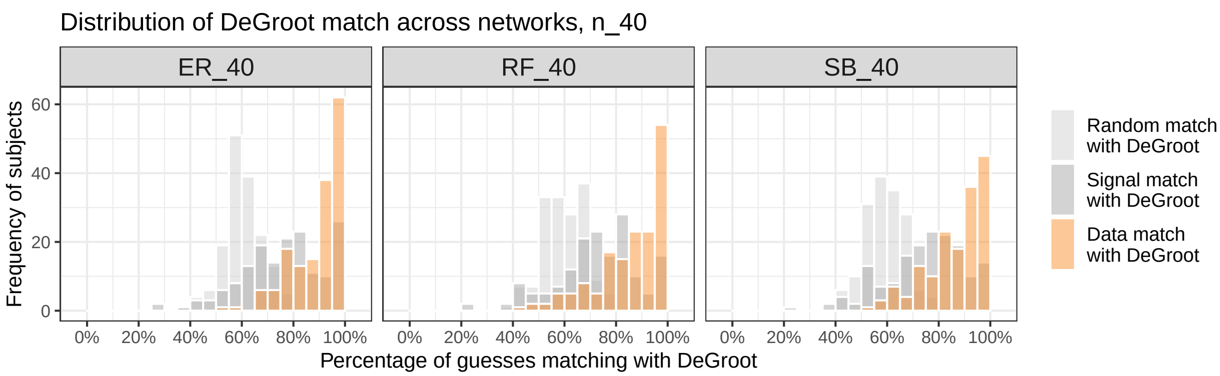


(b) DeGroot simulation with 15% trembling

Figure 17: Consensus achieved under DeGroot simulation with trembling.



(a) Fraction of guesses that match with DeGroot, compared to baselines.



(b) Fraction of subjects that match with DeGroot, compared to baselines

Figure 18: Percentage of guesses/subjects match with the DeGroot rule. Guesses matching DeGroot prediction are in orange; (Simulation) Guessing signal matching DeGroot prediction are in dark grey; (Simulation) Guessing randomly matching DeGroot prediction are in light grey. (a) Roughly 88% of guesses match with DeGroot predictions, significantly higher than the other two baselines of 75% and 60% respectively. (n=46,080: 11,520 per network) (b) 80% of subjects in ER match with DeGroot predictions at least 80% of the time; these fractions are 72% in the RF and 76% in the SB. This is again compared to the baseline of how well guessing signal matches with DeGroot predictions: Only 44% of pseudo subjects' guesses (if guessing only signal) in ER match with DeGroot predictions at least 80% of the time (37% in RF, and 41% in SB); A negligible fraction of pseudo subjects' guesses (if guessing randomly) match with DeGroot predictions at least 80% of the time. (n=960: 240 per network)

Table 12: Fraction of guesses imitate leader against DeGroot prediction

	Correctly follow leader			
	OLS (Bayesian predicts 1)	Logit	OLS	OLS
(Intercept)	0.10^{***}	-2.20^{***}	0.18^{***}	0.18^{***}
	(0.02)	(0.18)	(0.03)	(0.02)
RF_40	-0.06^{***}	-0.91^{**}		
	(0.02)	(0.39)		
SB_40	0.04^{**}	0.41^{*}		
	(0.02)	(0.21)		
period			-0.01^{***}	
			(0.00)	
round				-0.02^{***}
				(0.01)
R^2	0.01		0.01	0.01
Adj. R^2	0.01		0.00	0.01
Num. obs.	1870	1870	1870	1870
AIC		1388.23		
BIC		1404.83		
Log Likelihood		−691.12		

$^{***}p < 0.01$; $^{**}p < 0.05$;$^{*}p < 0.1$

Table 13: Fraction of guesses following signal against DeGroot prediction

	Always follow signal			
	OLS (Stubbornness predicts 1)	Logit	OLS	OLS
(Intercept)	0.25^{***}	-1.07^{***}	0.35^{***}	0.40^{***}
	(0.01)	(0.07)	(0.02)	(0.02)
RF_40	0.04	0.21		
	(0.04)	(0.21)		
SB_40	0.04^{*}	0.22^{*}		
	(0.02)	(0.12)		
period			-0.01^{***}	
			(0.00)	
round				-0.03^{***}
				(0.01)
R^2	0.00		0.00	0.01
Adj. R^2	0.00		0.00	0.01
Num. obs.	9366	9366	9366	9366
AIC		11185.38		
BIC		11206.81		
Log Likelihood		−5589.69		

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

C Related experiments

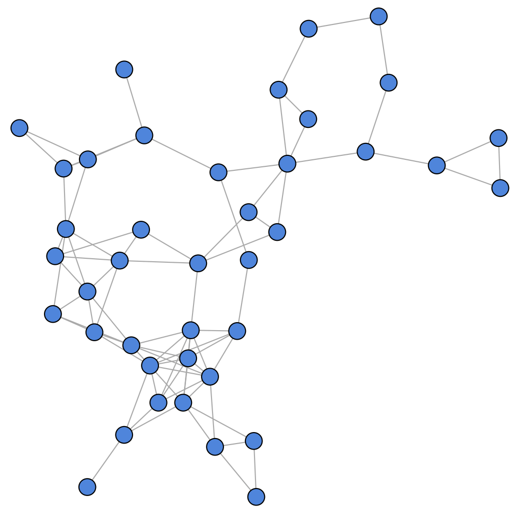


Figure 19: RGG graph of n=40 with average outdegree 4

In a recent paper, Chandrasekhar et al. (2020) looked at the mixture model of Random Geometric Graphs and Erdös-Rényi Graphs. We denote it as the RGG network from this point forward. This model captures the idea of sparse and clustered networks from the real world where the share of 'clans' — a set of nodes that are more connected among themselves than to those outside — is non-vanishing as n grows. This feature of inward-looking clans is also present in the 5-player communities within the Stochastic Block network. Under DeGroot updating rule, 'clans' being inward-looking facilitates the breakdown of consensus.

The network generation process is as follows: These exists a Poisson point process on the latent space $\Omega = [0,1]^2 \subset \mathbb{R}^2$, which determines the latent location of n nodes, with uniform intensity $\lambda > 0$. For any subset $A \subset \Omega$, $n_A \sim Poisson(\nu_a)$, where $\nu_a := \lambda \int_A dy$. If the Euclidean distance between two nodes i and j are at most $r = 0.2$, then i and j are linked with probability $\alpha = 0.95$. Otherwise, they are linked with probability $\beta = \alpha/(3n) < \alpha$. These parameter specifications were selected to ensure strong connectedness in the networks generated. Figure 19 presents an example of the RGG network which is also used in the experiment.

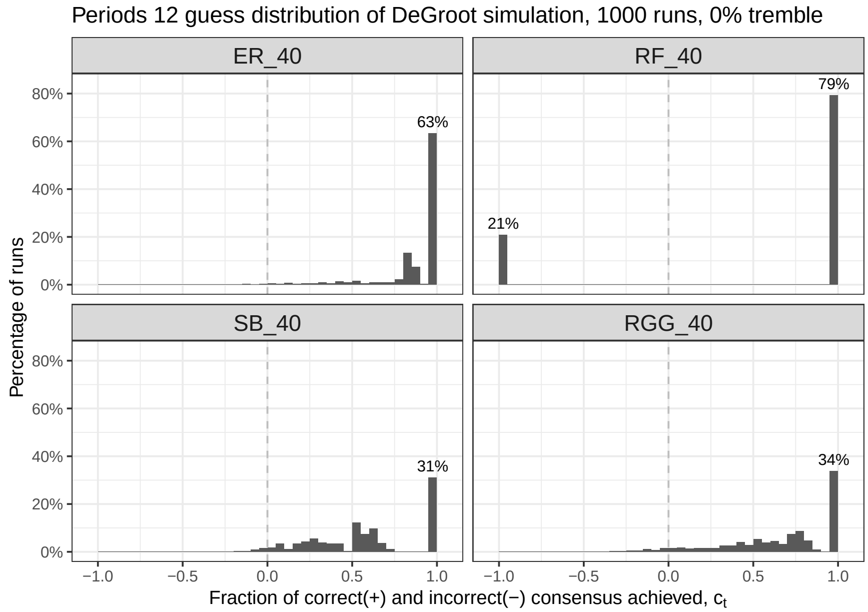
Figure 20a presents the simulation results of DeGroot updating rule on the RGG network and compares it with the Erdös-Rényi, Royal Family and Stochastic Block network. These simulations suggest that the RGG "lies between" the Erdös-Rényi and Stochastic Block network. The quartiles and the mean of the distribution of simulated c_t confirm this (Table 14). We observe the same results in the experiment (Figure 20b).

The Erdös-Rényi and Stochastic Block networks are canonical networks. Given the simulations and the experimental findings noted above, for expositional reasons, we felt it was best to present the Erdös-Rényi and Stochastic Block networks in the main text and move the RGG network to

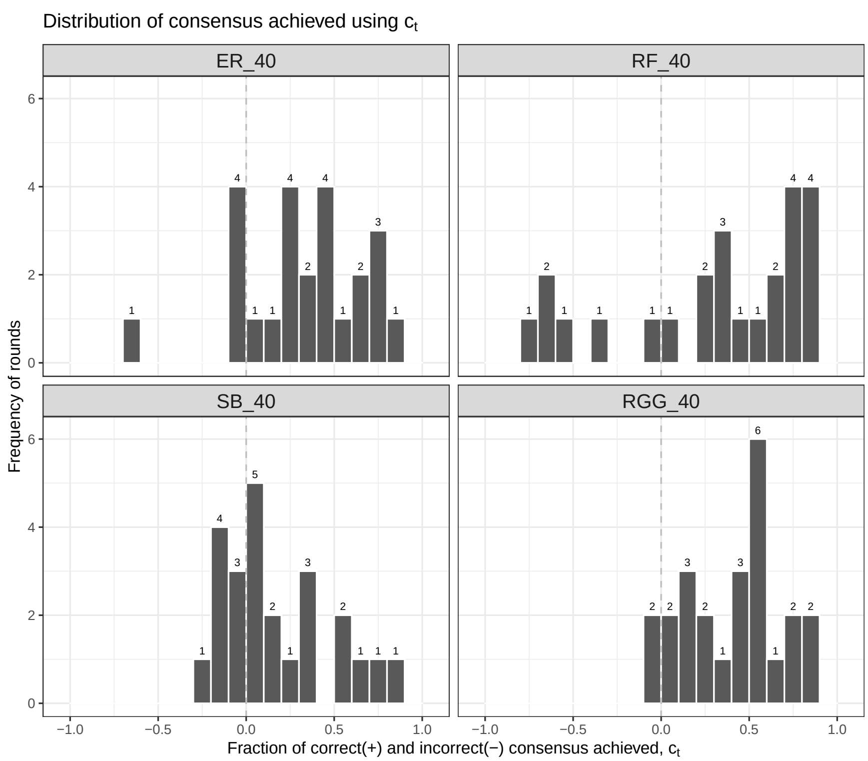
Table 14: Quartile and Mean of c_t under DeGroot simulation.

type	1st quartile	2nd quartile	3rd quartile	mean
RF	1	1	1	0.792
ER	0.95	1	1	0.953
RGG	0.825	0.925	1	0.882
SB	0.75	0.875	1	0.864

this Appendix.



(a) Distribution of c_t under Simulation (with RGG)



(b) Distribution of averaged c_t from experimental data (with RGG)

Figure 20: Distribution of averaged c_t. (a) The simulation of 1000 sets of signals shows that the distribution of consensus achieved by RGG lies between ER and SB. (b) Our experiment confirms the results from the simulation. (n=24 per network)

D Experimental Design

The experiment took place at the Laboratory for Research in Experimental and Behavioral Economics (LINEEX) at the University of Valencia. Subjects were recruited through the online recruitment system of LINEEX. All subjects who participated in this study provided informed consent at the LINEEX laboratory, and the procedure of this study was approved by the Institutional Review Board of the University of Valencia. In the experiment, subjects interacted through computer terminals in the LINEEX laboratory, and the experimental software was programmed in HTML, PHP, Javascript, and SQL.

Upon starting an experimental session, subjects read the paper-based instructions, which were also read aloud by an experimenter to guarantee that everyone received the same information (Supplementary Materials). The subjects were then provided with a step-by-step interactive tutorial on their computer screen, which allowed them to get familiarized with the software interface and the game (Figure 21). To clarify possible consequences of guesses in different periods of a round, subjects were shown a sample network (with only 10 players but with similar features as the network used in the actual game, depending on the experimental condition) highlighting what guesses would be observed by subjects as a decision maker from their neighbours, and their neighbours' neighbours.

Details about the decision screens were also provided to subjects: during any period of the game, each subject was shown the colour of the ball initially drawn, and guesses made by neighbours in the network during the previous period (Figure 23). Subjects also had the ability to view guesses made by those individuals (and themselves) in earlier periods of the game through a slider button. At the end of a round, a feedback screen revealed information about the payoff effective period that has been randomly selected, the guesses made by the subject and all others in thin period, the bag actually selected, and consequently the payoffs received by the subject in this round (0 or 3 euros depending on whether the guess matches the bag) (Figure 24). Prior to starting the first round of the game, all subjects also filled up a short questionnaire (4 questions) about their comprehension of the decision screens (Figure 22). Correct answers were shown after each guess made by the subjects.

To prevent long inactivity during the game, subjects were asked to make all guesses within 30 seconds (in any period of any round). If no guess was made before this time limit, a guess was made automatically, replicating the most recent guess or choosing at random in the first period. Throughout the experiment, all guesses, with no exception, were made by subjects within this time limit.

E Raw data

Figures 25 to 27 present the evolution of the average guesses of each network treatment (ER, SB, RF), group (1-4), and round (1-6) from the experiment.

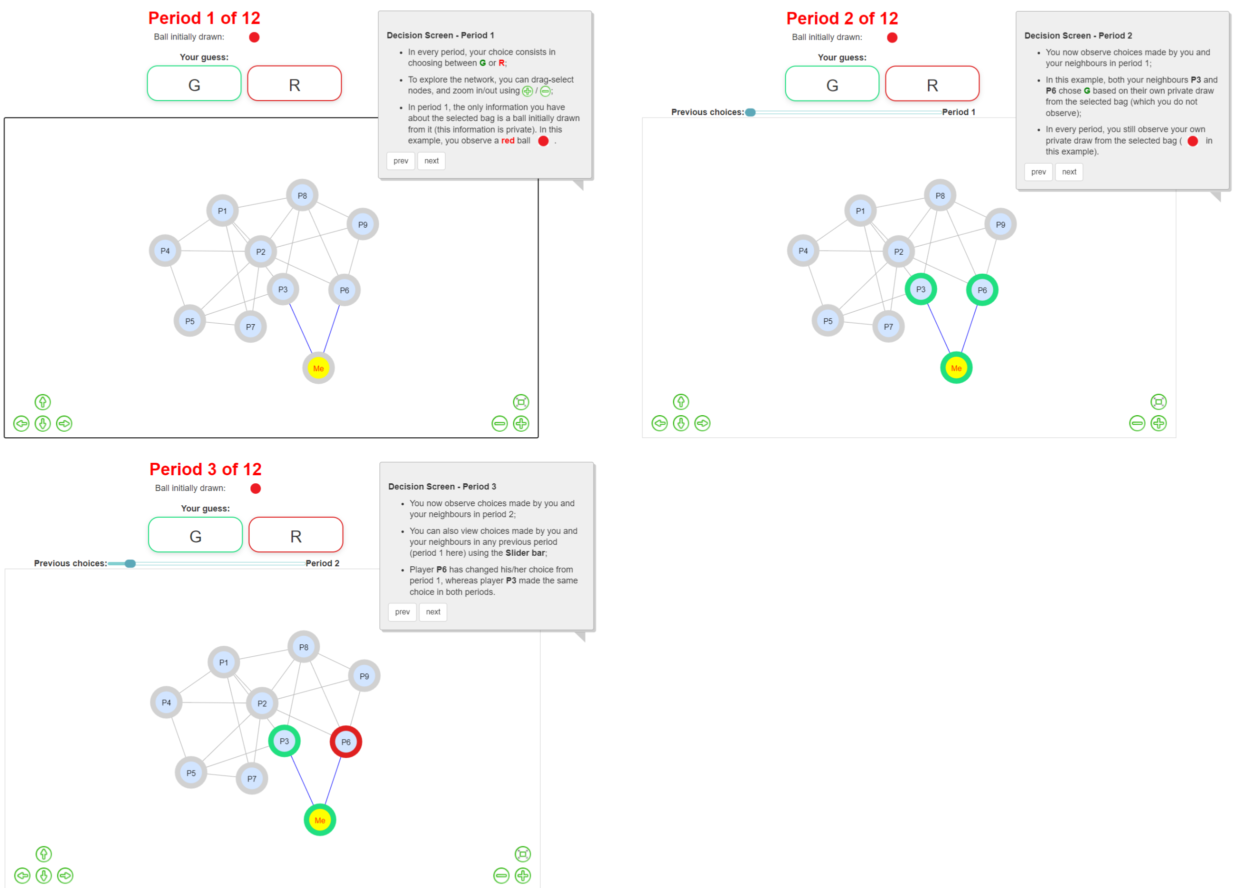


Figure 21: Tutorials from the experiment

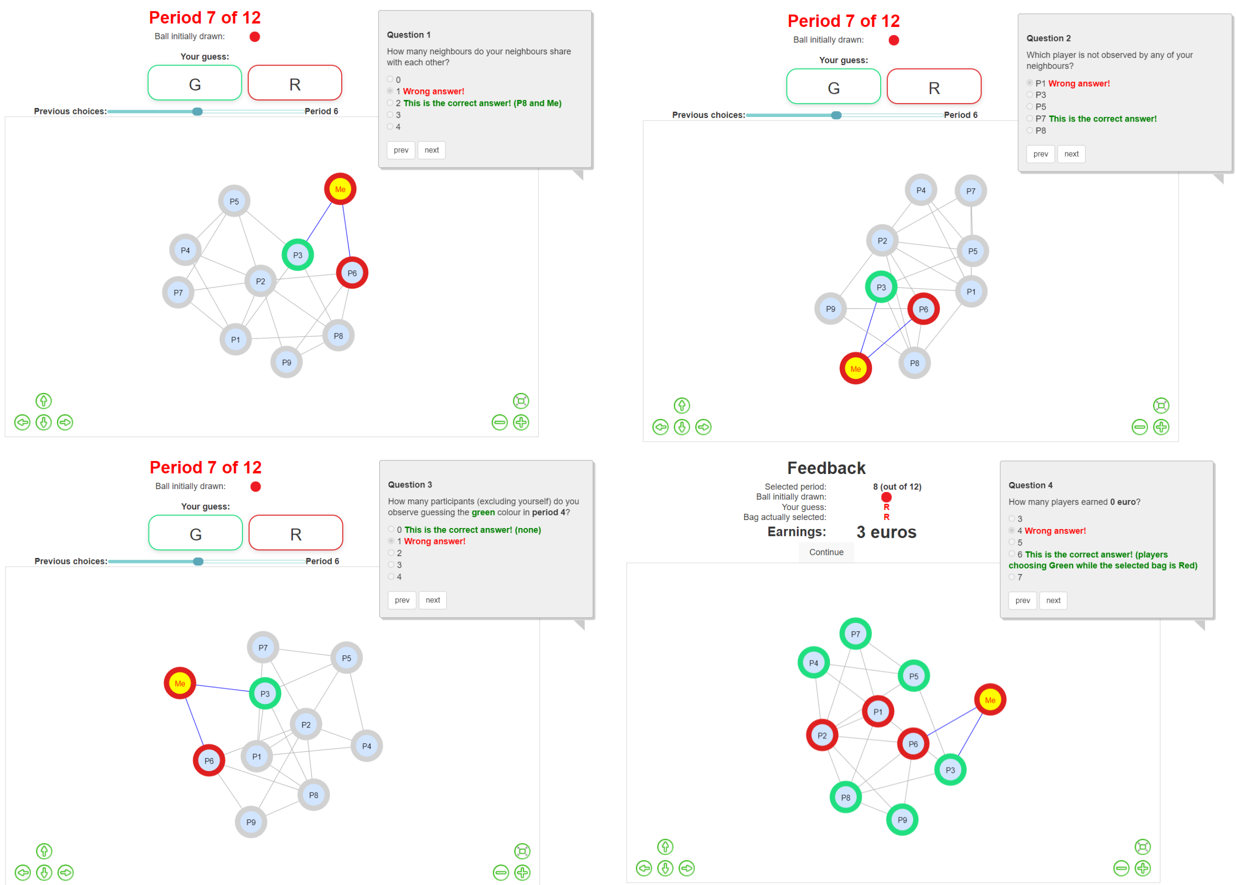


Figure 22: Questionnaires from the experiment

Figure 23: Screenshots from the experiment during the game

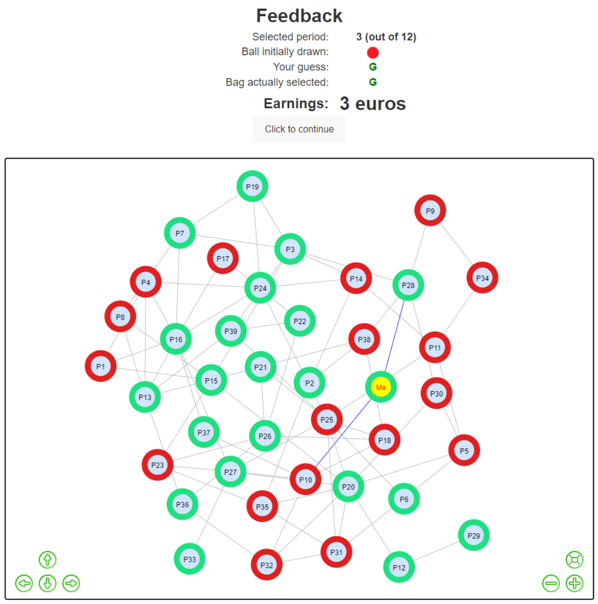


Figure 24: Feedback screen from the experiment during the game

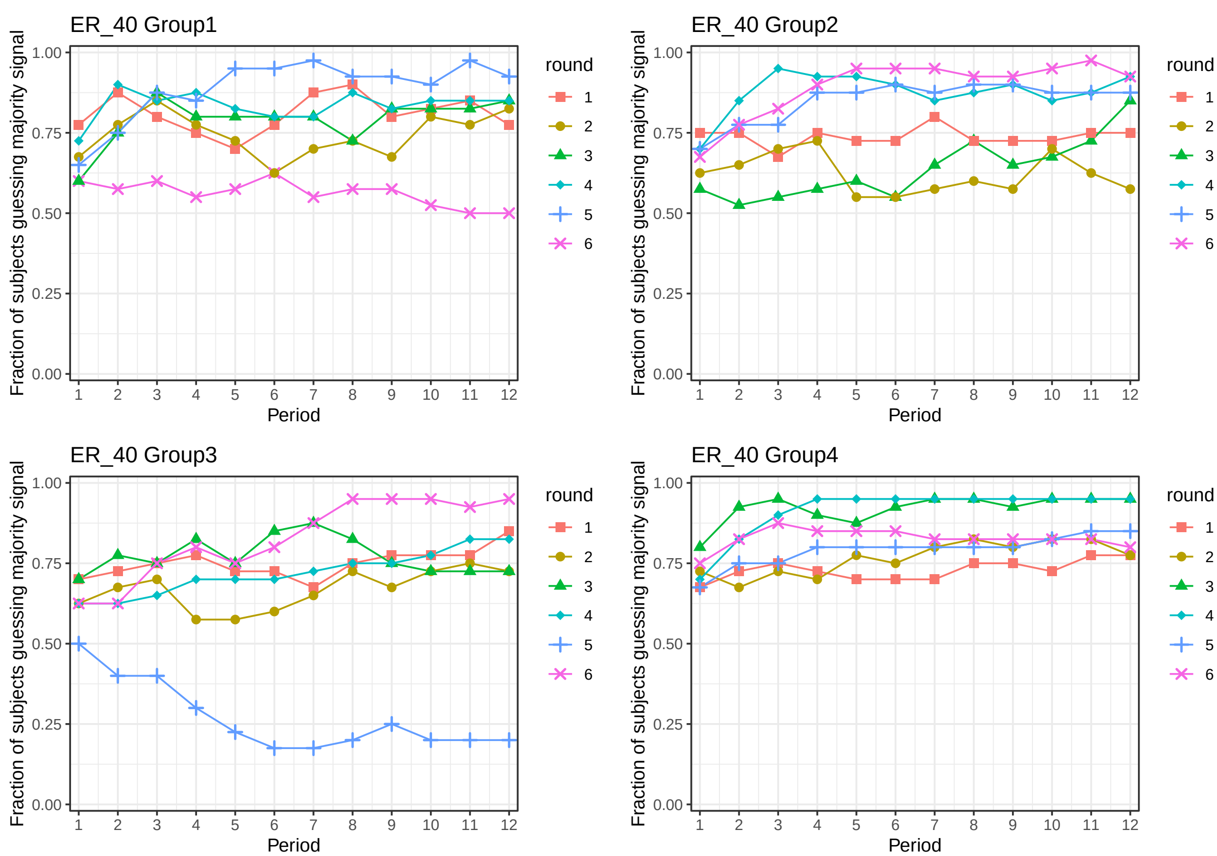


Figure 25: Experimental results — Development of guesses n=40

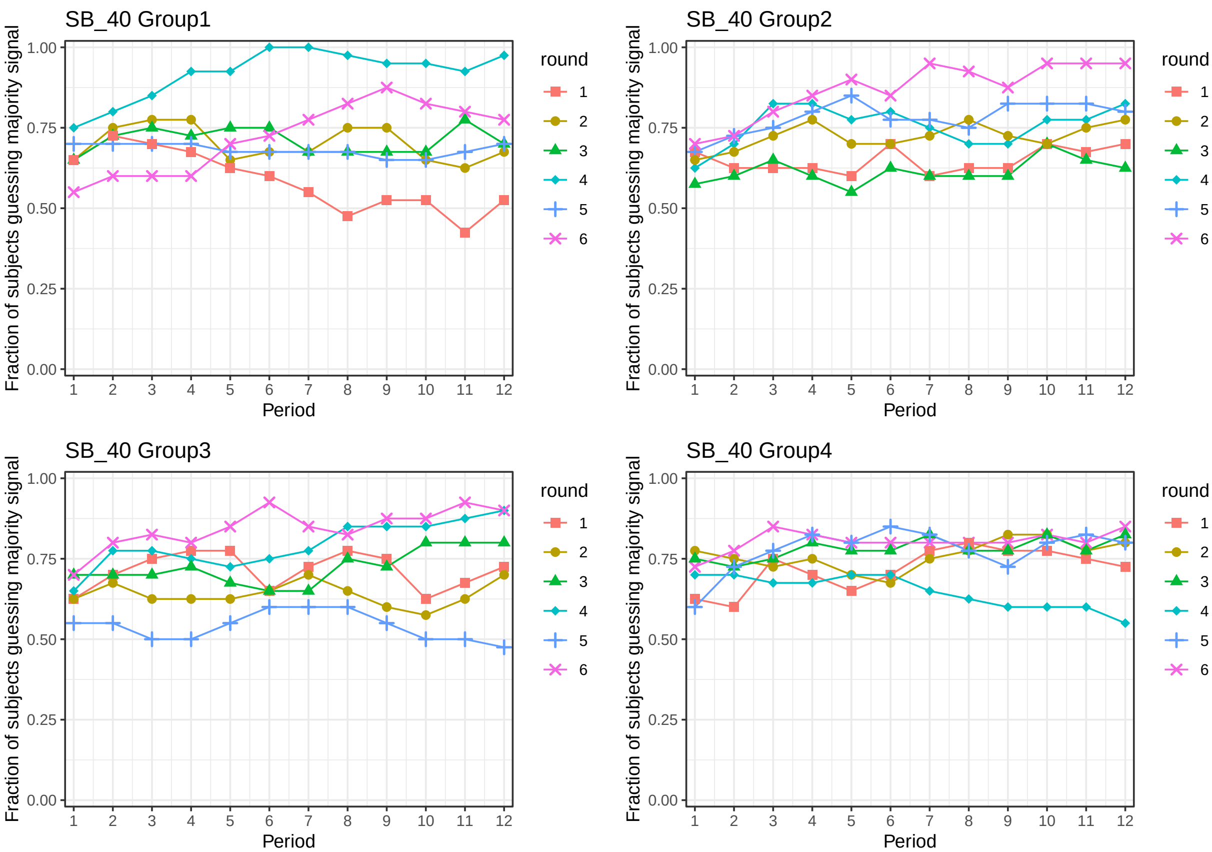


Figure 26: Experimental results — Development of guesses n=40

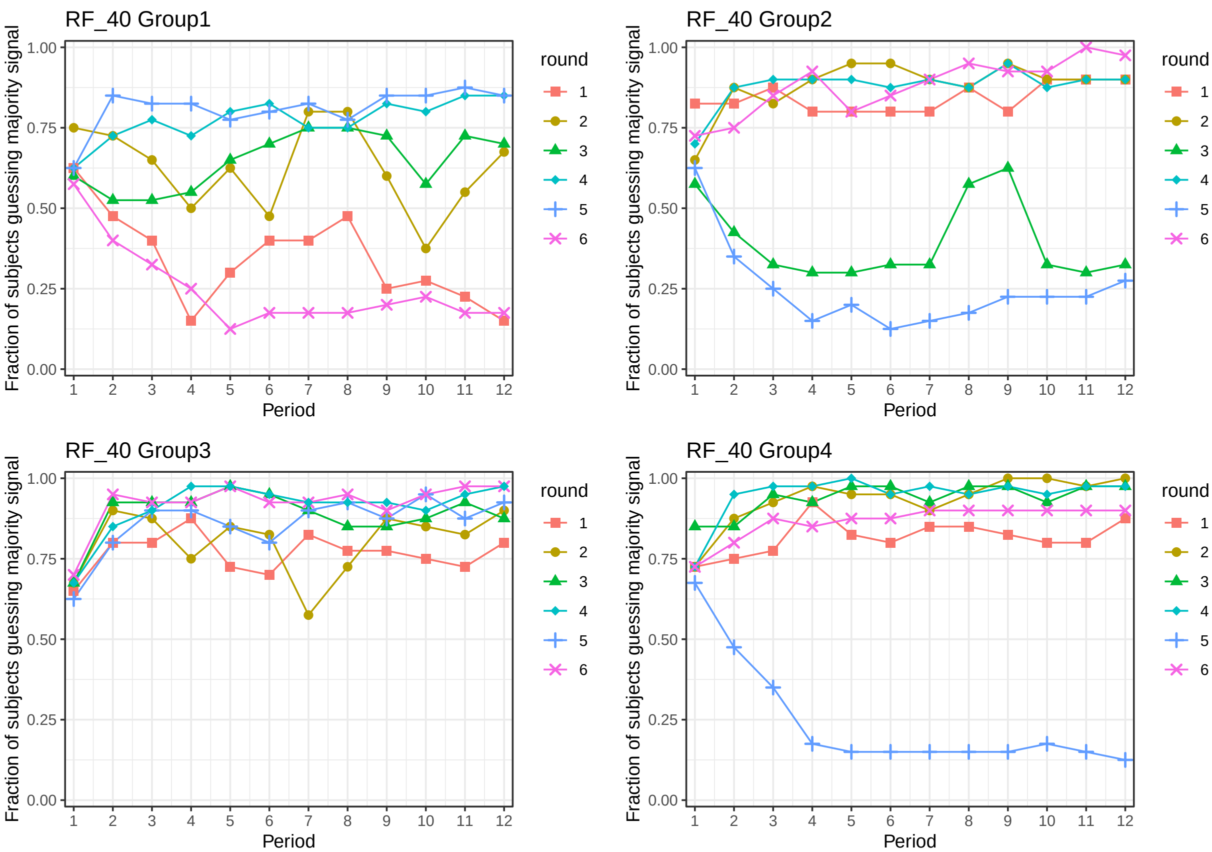


Figure 27: Experimental results — Development of guesses n=40