CENTRE FOR DECISION RESEARCH & EXPERIMENTAL ECONOMICS

Discussion Paper No. 2024-07

Ying Chen, Tom Lane and Stuart McDonald

October 2024

Social Identity in Network Formation

CeDEx Discussion Paper Series
ISSN 1749 - 3293

CENTRE FOR DECISION RESEARCH & EXPERIMENTAL ECONOMICS

The Centre for Decision Research and Experimental Economics was founded in 2000, and is based in the School of Economics at the University of Nottingham.

The focus for the Centre is research into individual and strategic decision-making using a combination of theoretical and experimental methods. On the theory side, members of the Centre investigate individual choice under uncertainty, cooperative and non-cooperative game theory, as well as theories of psychology, bounded rationality and evolutionary game theory. Members of the Centre have applied experimental methods in the fields of public economics, individual choice under risk and uncertainty, strategic interaction, and the performance of auctions, markets and other economic institutions. Much of the Centre's research involves collaborative projects with researchers from other departments in the UK and overseas.

Please visit http://www.nottingham.ac.uk/cedex for more information about the Centre or contact

Samantha Stapleford-Allen
Centre for Decision Research and Experimental Economics
School of Economics
University of Nottingham
University Park
Nottingham
NG7 2RD
Tel: +44 (0)115 74 86214
Samantha.Stapleford-Allen@nottingham.ac.uk

The full list of CeDEx Discussion Papers is available at

http://www.nottingham.ac.uk/cedex/publications/discussion-papers/index.aspx

Social Identity in Network Formation[*]

Ying Chen, Tom Lane and Stuart McDonald

This version: October 10, 2024

Abstract

Using a laboratory experiment, we study the evolution of economic networks in the context of fragmented social identity. We create societies in which members can initiate and delete links to others, and then earn payoffs from a public goods game played within their network. We manipulate whether the society initially consists of segregated or integrated identity groups, and vary whether societal mobility is high or low. Results show in-group favouritism in network formation. The effects of original network structure are long-lasting, with initially segregated societies permanently exhibiting more homophilic networks than initially integrated ones. Moreover, allowing greater social mobility results in networks becoming less rather than more integrated. This occurs in part because eviction from networks is based on out-group hostility when societal mobility is high, and on punishing free riders when mobility across groups is low.

Keywords: Social identity; Social network; In-group bias; Homophily; Laboratory experiments.
JEL classifications: C92; D85; D91; Z13.

---

[*] **Acknowledgements:** This research is sponsored by the Centre for Decision Research and Experimental Economics (CeDEx China) at the University of Nottingham Ningbo China. We have benefited from valuable comments by Xueheng Li, and seminar participants at Newcastle University, Masaryk University, Xi'an Jiaotong-Liverpool University, Sun Yat-sen University, Southern University of Science and Technology, ESA 2024 European Meeting (Helsinki). Ying Chen acknowledges support from Di Wang for hosting half of the experiment sessions at the China Center for Behavioural Economics and Finance (CCBEF), Southwest University of Finance and Economics, Chengdu, China. The research protocol was approved by the University of Nottingham Research Ethics Subcommittee on December 9, 2022. The experiment is pre-registered at AEA RCT Registry, named "How social identity affect network formation and public good provision?", by Ying Chen on February 13, 2023, and can be retrieved from https://doi.org/10.1257/rct.10922-1.0
Ying Chen: Research Fellow, Centre for Decision Research and Experimental Economics (CeDEx China); School of Economics, University of Nottingham Ningbo China. Email: yingchen.econ@outlook.com.
Tom Lane: Senior Lecturer, Newcastle University Business School. Email: tom.lane@newcastle.ac.uk.
Stuart McDonald: Associate Professor, School of Economics, University of Nottingham Ningbo China. Email: stuart.mcdonald@nottingham.edu.cn.

1. Introduction

Networks can have important economic and social consequences, primarily because economic and social benefits can be distributed via network connections. Consequently, individual well-being can be impacted by decisions related to social inclusion and exclusion that are made in social networks. Inclusion in a network may bring both material advantages and disadvantages depending on the nature of individuals invited into the social network. Exclusion from networks may lead to the deprivation from economic and social opportunities, when an individual derives economic and social value from its membership. In short, the dynamics of network connections are pivotal in determining individuals' economic success within a society (Jackson, 2008).

Sociologists have long argued – and more recently economists have joined them in doing so – that when individuals make decisions about whom to connect with, the networks they form usually contain the special segregating feature of *homophily*, wherein individuals sharing similar characteristics are more likely to link with each other (e.g., Currarini et al., 2009; McPherson, et al., 2001). This is important, since if network membership is based upon shared social identity, those belonging to out-groups that are excluded from profitable social networks, are more likely to end up among society's losers (Small & Pager, 2020). This is one reason for why social inclusion is a key platform of the modern movement to promote Diversity, Equity and *Inclusion* (Barnett, 2020).

Numerous economic studies have identified the tendency for homophily in economic networks (see, e.g., Barr et al, 2012; Currarini & Mengel, 2016; Gallo et al., 2022). However, how such preferences, when allowed to manifest and mutate over many repeated interactions, translate into the dynamic evolution of network structures, has not been studied thoroughly. One question is the extent to which in-group favouritism influences the linking decisions in forming networks. The other question is whether in-group favouritism persists in the scenario when individuals are asked to maintain connections with others, particularly when they have the chance to observe the cooperative behaviour of those they are already connected with. In a dynamic network setting, it remains to be studied whether social identity will diminish in significance over time, eventually leading a breakdown in segregation.[1]

[1] Another reason for studying homophily in a dynamic setting is that baseline preferences may not aggregate in a straightforward way. Consider, for instance, the classic segregation model of Schelling (1969), with its striking conclusion that even very mild preferences for interacting with people from the same group will lead over time to complete segregation.

In this paper, we investigate the role of social identity in endogenous network formation in the context of a multiple-round public goods game. First, we are interested in understanding whether subjects are more likely to build and maintain connections with others who have the same social identity. Hence, we want to understand whether there is an economic foundation to the phenomenon of homophily in social networks. We conduct a laboratory experiment in which subjects, who belong to one of two kinds of artificial social identities, repeatedly play a congestible public goods game with the neighbours in their network. A key feature of this experiment is that subjects can adjust before each round of play by connecting with new network members and disconnecting from existing network members, thus causing the network structure to change in each round of the game. Forming a new pair of neighbours requires mutual agreement by both parties while breaking an existing relationship only needs action by either of the two parties. When deciding whether to form a new pair, each subject can observe each other's social identity; and in making decisions over whether to maintain an existing relationship, each subject can observe each other's social identity, as well as her previous performance. Each subject's payoff depends on two factors. First, the payoff depends on the combination of her own choice of contribution plan and the most common choice of contribution plan of all neighbours in her network. Second, it depends on the size of her network: networks which are either too large or too small are inefficient. Therefore, payoff maximisation would result from selecting an optimal number of trustworthy neighbours.

We also study the process under which social networks evolve to gain an understanding of how different social contexts influence network formation and homophily. We explore this using a 2×2 factorial design. The first design dimension is *Mobility*, which is defined as the probability that any given relationship is allowed to be changed at a given moment in time. In half of the treatments, this probability is high, as each pair of subjects has an 80% chance of being able to change their relationship from strangers to neighbours, or vice versa, before undertaking a round of the congestible public goods game. In the other half of the treatments, the probability of admitting outsiders or deleting insiders is only 30%. We use this dimension to model the stability of networks within a society. An example of a *Low Mobility* society could be a rural society, with infrequent population inflows, limiting the opportunities to connect or disconnect with economic partners from other villages. An example of a *High Mobility* society would be a metropolis, in which it is easy to meet new contacts or disassociate with existing ones. In short, the Mobility dimension induces a churning effect into the network, allowing members the option of updating acquaintances with new friends in the group.

The second dimension is *Integration*, which is defined as the composition of in-group and out-group neighbours in the initial social network. In half of the treatments, subjects will start the game with an assigned social network in which all neighbours are of the same social identity as themselves. This is akin to a society in which historical forces have resulted in, for instance, high levels of racial segregation. In the other half of the treatments, all neighbours in the assigned social network for each subject come from the opposite social identity group. This resembles, for instance, a society with a high initial level of mixing between racial groups. We use this dimension to explore whether the initial structure of a society affects the composition, and associated outcomes, that its networks eventually gravitate towards.

Our experimental results show subjects care about social identity when forming networks, even though the social identity we used in the lab is artificially generated. Subjects have a strong preference to invite those who have the same social identities as themselves to join their social networks. This in-group favouritism appears in all treatments but decreases when the game is played repeatedly.

Second, we find that higher social mobility increases network homophily, because subjects tend to delete social identity out-group members from their networks under *High Mobility* but not under *Low Mobility*. Although we hypothesised that a more fluid society would promote inter-group interactions and reduce in-group bias, our experiment shows it often leads individuals to reinforce their in-group networks. While this may strike the reader as a counterintuitive result, it supports recent – and perhaps also surprising – real-wrld evidence that socioeconomic segregation is much more severe in larger cities than in smaller ones (Nilforoshan, 2023). It suggests that simply creating opportunities for greater mobility in society does not necessarily promote integration and may even backfire. On the contrary, in the *Low Mobility* treatments, the importance of social identity decreases – participants tend to eliminate free riders (irrespective of social identity) from their networks, possibly because they fear being stuck with these individuals in endless interactions.

Third, we find that the initial structure of the social network impacts the network structure in long-term. Initially more segregated networks are likely to end up with higher degree of network homophily. This finding highlights the intense difficulties facing policymakers aiming to combat homophily in social networks within societies the initial configuration of which they are unable to control. However, it also offers an important lesson for organisational design. When bringing together a new cohort of individuals (for instance, a new work unit, or a new intake of students), the network connections made in early interactions can be pivotal in the long term; if a cohort is drawn from diverse backgrounds, and the organisation hopes to avoid it segregating along identity lines, deliberately placing members into socially mixed project teams on day one may ultimately prove a fruitful intervention.

In summary, our experiments offer valuable guidance for managing teams, designing organizational structure, and creating policies aimed at promoting integration in a diverse society. The rest of our paper is organized as follows. Section 2 links this paper to existing literature. Section 3 presents the design of the experiment. Section 4 reports the main results and Section 5 concludes the paper.

2. Literature review

Our study contributes to the discussion of the effect of social identity on human behaviour. Previously the experimental economics literature has shown that people exhibit in-group favouritism in the sense that they have a preference towards linking with individuals who possess the same social identity as themselves (Balliet et al, 2014; Li, 2020). As examples, Charness et al. (2007) show that sharing the same social identity increases the cooperation decisions in the Prisoner's Dilemma game and the defection decisions in the Battles of Sexes game; Bernhard et al. (2006) find that third-party punishers spend less money to punish norm violators who share the same social identity as the punisher; Hargreaves Heap et al. (2009) find that investors give more money to group members than outsiders in a trust game; Chen & Li (2009) find that people distribute more money to in-group members than out-group members; Tan & Bolle (2007) find that sharing the same social identity increases cooperation in the linear public good game. The studies summarized above provide a corpus of evidence that participants treat in-group members and out-group members differently in terms of material outcomes.

Our study also relates to several laboratory experimental studies that document homophily in their findings.[2] For instance, Currarini & Mengel (2016) find that when participants are given the freedom to choose in- and out-group matching for playing games together, they are more likely to express a positive

[2] Some recent theoretical studies on networks have started to focus on homophily as well. Currarini et al. (2009) is a seminal paper in which they model the process of friendship network formation in American high schools and explain that a preference of same-type friendship leads the members of large demographic groups to earn more links in a network, thus demonstrating the emergence of homophily. Currarini et al. (2016) add to this in showing that the preference for same-type connections can be due to people having different chances of meeting in- and out-group members. Bramoullé et al. (2012) studies the evolution of segregation in homophilic networks with an application to a dynamic network of scientific citations. More closely related to our study, another strand of literature studies the emergence of homophily in the formation of networks that contain some strategic interactions. Studies include Baccara & Yariv (2013) and Allmis & Merlino (2023) in which agents need to develop networks and the aim of it is to strategically invest into a public project with their network participants to maximise payoffs.

willingness to pay for the in-group match; Charness et al. (2014) find that when subjects are given both a group identity and a monetary identity in playing the public goods game, their monetary identity becomes more salient in creating homophily; Friebel et al. (2021) and Mengel (2020) find that both men and women are more likely to develop connections with same-gender group members in their networking activities.[3]

Our study also relates to the literature about public good provision in endogenously formed networks. These studies focus on the effect of different means of network formation on the level of contribution to the public good (Ehrhart & Keser, 1999; Ahn et al., 2008; Ahn et al., 2009; Charness & Yang, 2014; Coricelli et al., 2004). These studies support the conclusion that an endogenously formed group improves the level of cooperation. Among those papers, Ahn et al. (2009), similar to our paper, use a congestible public goods game. However, these studies have not explored questions relating to social identity.

Our interest in the effects of mobility on network evolution is shared with three experimental studies: Gallo et al. (2022), Riyanto & Teh (2020) and Rand et al. (2011). The three papers focus on the impact of the rate of change in the social environment on the level of cooperation. In their network formation processes, they alter the rate of network updating, allowing either zero, a small, or a large fraction of pairs of relationships to be updated in each round of a cooperation game. The three papers all document that in a fast-change environment, the level of contribution is higher because subjects will form links with cooperators and break links with defectors. In Riyanto & Teh (2020), they further find that whether the flexibility of changing the network relationships could facilitate cooperation depends on the structure of the network at the beginning of the experiment. This design shares similarities with our Mobility manipulation. Our study differs from these both by focusing on public goods contributions and by introducing social identity.

There are some experiments that discuss the effect of the initial network structure on individual decisions and their outcomes. Rosenkranz & Weitzel (2012) find that the level of cooperation in the local public goods game depends on the structure of the network assigned to participants. They show that participants are more likely to cooperate in star networks or when network structure is complete (i.e., when all network members are connected). Van Leeuwen et al. (2019) find that the central position in a network can play a

[3] Homophily is also identified in empirical studies. Banerjee et al. (2013) and Jackson (2014) find that the networks for spreading a microfinance product in Indian villages show homophily by participants' caste designations. Besides that, there is a strand of empirical literature documenting evidence of gender-based network homophily in referral networks (Zelter, 2020), in citation networks (Davies, 2022; Ductor & Prummer, 2024) and in friendship networks (Jackson et al., 2023).

key role in maintaining a high level of public good provision in the presence of endogenous exclusion of network members. Riyanto & Teh (2020) compare the efficiency of cooperation behaviour when assigning a star network or a circle network as the initial group structure. They find that the efficiency is higher in a star network because this structure contains a central agent, when compared to circle networks.

3. Experimental design

We study the behaviour of a small "society" in the laboratory. Each "society" has 12 participants. They are evenly divided into two social identity groups, and each of them is assigned a unique ID showing which of the two social identity groups she belongs to. Our procedures of dividing participants into social identity groups and assigning them unique IDs are discussed in Section 3.1. Then, the 12 participants are engaged to perform 25 rounds of the congestible public goods game, which is introduced in Section 3.2. Section 3.3 discusses treatment variations and Section 3.4 documents procedures.

3.1 Social identity assignment and enhancement

To gain their social identities, each participant obtains a unique ID showing which of the two social identity groups they belonged to. To achieve this, each participant was asked to complete a music task first and then a discussion task. The two tasks follow the minimal group paradigm in Chen & Li (2009), in which an initial task divides participants into either of the two artificial social identity groups and the discussion task serves the purpose of potentially enhancing these newly created identities.

Music task – Participants were first asked to play a music task aimed at inducing the heterogeneous social identities in the lab. In this task, all 12 participants reported their preferences over five pairs of classical music excerpts, each of which contained one excerpt from Variations Goldberg (henceforth, VG) and one excerpt from The Well-Tempered Clavier, Book II (henceforth, WT).[4] After listening to a pair of excerpts, they needed to allocate nine points between the two, giving more points to the excerpt they preferred. Then they were assigned, in equal numbers, into either the VG or WT group. This was done by ranking the participants according to the total points they allocated to VG, placing those ranked from 1 to 6 in the VG group and those ranked from 7 to 12 in the WT group.

---

[4] To ensure trivial (unimportant) differences within each pair, we chose two compositions written by Bach and performed by the same pianist. All the excerpts were played by a stereo loudly in the laboratory.

Discussion task – Next, to enhance social identity, we gave each music group ten minutes to have a discussion task. They talked by typing and sending messages to all group members but without face-to-face interaction. In these ten minutes, they needed to solve five questions together. To motivate the group members to talk and exchange their opinions sufficiently, the five questions were either open questions related to the music they had just heard, or questions related to their daily routines. At the end of the discussion, they needed to report answers to the five questions independently to the computer. Since this chatting period was solely for social identity enhancement, we did not incentivise the task.

At the end of the two tasks, each participant was given a unique ID in the format of VG (or WT) plus a number from 1 to 6. This unique ID was revealed in stages. First, at the end of the music task, each participant knew only that she was assigned to either the VG or WT group. Next, during the discussion task, to distinguish between senders of different messages, the six participants of the same group used temporary IDs to conduct their dialogue. Finally, at the end of the discussion task, after they had submitted their answers to the computer, participants were given their unique IDs, which they were told they would keep for the rest of the experiment. The temporary IDs in the discussion task could not be matched to the IDs participants were eventually assigned. This meant that, while the discussion task could enhance social identity, it would not lead to participants developing relationships with specific group members that could influence their decisions in the subsequent incentivised task.

3.2 The congestible public goods game

After their social identities have been allocated and their unique IDs are obtained, the 12 participants played 25 rounds of the congestible public goods game. Each round contained two steps. In Step 1, they were given opportunities to select neighbours and delete existing neighbours from their network.[5] Then in Step 2, they were matched with their neighbours determined in Step 1 and made contribution decisions. Each participant's payoff in each round depends on (1) how many neighbours they have and (2) the contribution decisions of herself and her neighbours.

In this game, being neighbours is a symmetric relationship, i.e., having someone as a neighbour also entails being their neighbour. In each round, a pair of participants would be defined as neighbours either if this was the pre-existing arrangement and neither acted to end it, or if they acted in Step 1 to become new

[5] The exception was the first round, in which subjects' neighbours were automatically assigned to them and they progressed straight to Step 2. See details in Section 3.3.

neighbours. This would happen if both participants in the pair chose to initiate a link to the other. If only one of the two participants chose to initiate such a link (or neither did), they would not become neighbours.

A pre-existing pair of neighbours would be broken if, in Step 1, either of the two participants took the unilateral action to cancel the link. The 12 participants made these linking decisions in Step 1 simultaneously. There was no fee for either initiating or deleting a link. At the end of this step, the computer programme informed each participant how many neighbours she had, and who they were. Each participant only knew about their own neighbours; they were not informed which of the other participants were neighbours with each other.

Note that we are interested in whether each participant prefers to have neighbours with in-group identity. To this purpose, in Step 1, when the two participants were considering whether to initiate links to each other, they could see each other's social identities and unique IDs. Other information, such as past contribution decisions, was not presented. This resembles the real world in which only surface level characteristics are visible when first encountering a new person. However, when a pair of pre-existing neighbours were considering breaking their relationship, besides social identity and unique ID, they could also see each other's contribution decision from the last round. This also resembles the real world in that one can observe not only surface level characteristics but also the actual behaviour of those in their existing network.

After knowing their neighbours, each participant then made contribution decisions in Step 2. Table 1 shows the game's payoff structure. Each participant has four available contribution plans: Plan 1, Plan 2, Plan 3 and Plan 4, of which she needs to choose one. We describe Plan 1 as "the highest plan" and Plan 4 as "the lowest plan" in the experimental instructions. When playing the game with a given number of neighbours, a participant's payoff depends on the combination of (a) her own chosen plan and (b) the most common chosen plan of her neighbours. For any given action by oneself, a player's payoff is greater the higher the most common plan of their neighbours. Meanwhile, for any given behaviour by one's neighbours, the player can increase their payoff by shifting from a higher plan to a lower one.

Second, a subject's payoff is associated with the number of neighbours they have. Table 1 indicates that, given the same outcome in terms of own and neighbours' choices, if the number of neighbours is between one and five then the payoff associated with this outcome weakly increases in the number of neighbours when the most common choice of neighbours is among Plans 1 to 3, and weakly decreases in the number of neighbours when the most common choice of neighbours is Plan 4; and, if the number of neighbours is between five and eleven, the payoff associated with this outcome weakly decreases in the number of neighbours. Therefore, the size of the network affects a subject's payoff depending on the collective behaviour of their chosen neighbours: the optimal network size is positive if the neighbours are contributing positively, while if all neighbours are free riders it would be better to be isolated.

Table 1: Earnings table.

Sub-table A. My number of neighbours = 0

I do not have to choose plans.
My earnings are 30 RMB.

Sub-table B. My number of neighbours = 1

My neighbours' most common Plan \ My Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	48.0	50.0	52.0	54.0
Plan 2	40.0	42.0	44.0	46.0
Plan 3	32.0	34.0	36.0	38.0
Plan 4	24.0	26.0	28.0	30.0

Sub-table C. My number of neighbours = 2

My neighbours' most common Plan \ My Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	60.0	63.3	66.7	70.0
Plan 2	46.7	50.0	53.3	56.7
Plan 3	33.3	36.7	40.0	43.3
Plan 4	20.0	23.3	26.7	30.0

Sub-table D. My number of neighbours = 3

My neighbours' most common Plan \ My Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	75.0	78.8	82.5	86.3
Plan 2	56.3	60.0	63.8	67.5
Plan 3	37.5	41.3	45.0	48.8
Plan 4	18.8	22.5	26.3	30.0

Sub-table E. My number of neighbours = 4

My neighbours' most common Plan \ My Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	93.0	96.8	100.6	104.4
Plan 2	68.2	72.0	75.8	79.6
Plan 3	43.4	47.2	51.0	54.8
Plan 4	18.6	22.4	26.2	30.0

Sub-table F. My number of neighbours = 5

My neighbours' most common Plan \ My Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	111.0	114.8	118.7	122.5
Plan 2	80.2	84.0	87.8	91.7
Plan 3	49.3	53.2	57.0	60.8
Plan 4	18.5	22.3	26.2	30.0

Sub-table G. My number of neighbours = 6

My neighbours' most common Plan \ My Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	93.0	98.6	104.1	109.7
Plan 2	66.4	72.0	77.6	83.1
Plan 3	39.9	45.4	51.0	56.6
Plan 4	13.3	18.9	24.4	30.0

Sub-table H. My number of neighbours = 7

My neighbours' most common Plan \ My Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	75.0	81.9	88.8	95.6
Plan 2	53.1	60.0	66.9	73.8
Plan 3	31.3	38.1	45.0	51.9
Plan 4	9.4	16.3	23.1	30.0

Sub-table I. My number of neighbours = 8

My neighbours' most common Plan \ My Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	60.0	67.8	75.6	83.3
Plan 2	42.2	50.0	57.8	65.6
Plan 3	24.4	32.2	40.0	47.8
Plan 4	6.7	14.4	22.2	30.0

Sub-table J. My number of neighbours = 9

My neighbours' most common Plan \ My Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	48.0	56.4	64.8	73.2
Plan 2	33.6	42.0	50.4	58.8
Plan 3	19.2	27.6	36.0	44.4
Plan 4	4.8	13.2	21.6	30.0

Sub-table K. My number of neighbours = 10

My neighbours' most common Plan \ My Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	42.0	50.7	59.5	68.2
Plan 2	29.3	38.0	46.7	55.5
Plan 3	16.5	25.3	34.0	42.7
Plan 4	3.8	12.5	21.3	30.0

Sub-table L. My number of neighbours = 11

My neighbours' most common Plan \ My Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	39.0	47.9	56.8	65.8
Plan 2	27.1	36.0	44.9	53.8
Plan 3	15.2	24.1	33.0	41.9
Plan 4	3.3	12.2	21.1	30.0

It can therefore be seen that, for a given number of neighbours, the payoff structure captures the standard social dilemma features of a public goods game. Each player could maximise her earnings in each round by choosing the lowest plan, but this would increase the likelihood of reducing the payoffs of one's neighbours[6]; if all players select the lowest plan, the resulting payoffs are Pareto inferior to the outcome that would result if all players select the highest plan. For ease of explanation to participants, decisions are framed as plans rather than contributions of a share of one's endowment.[7] However, the payoff structure is equivalent to a setting where each participant's payoff is derived from a voluntary contribution mechanism game played with her neighbours, where the marginal per capital returns are decreasing in the number of neighbours.[8] This structure is such that, holding constant the modal plan of neighbours, the optimal number of neighbours is five (as long as the neighbours' modal plan is not Plan 4).[9] However, a player's preferred number of neighbours may also depend on her belief about the likely behaviour of others; she may desire more or less neighbours than five, if she believes such a number will lead to better average behaviour among her neighbours. At the end of Step 2, each participant was informed of the choice of each of her neighbours, and therefore the most common choice among them, as well as her payoff in this round.

---

[6] Deviating to the lowest plan only probabilistically reduces neighbours' payoffs, since it may not affect the most common plan chosen by those in a given neighbour's network. This is a slight difference from a standard public goods game. In fact, in this game, the social optimal is achieved if a majority of players select the highest plan and a minority free ride. However, since decisions are made simultaneously, selecting the lowest plan can still clearly be regarded as a potentially selfish action and therefore, for the purposes of our paper's research questions, we do not believe this non-standard feature of the public goods game is of much importance.

[7] Payoff of each subject is therefore based on neighbours' modal plan rather than mean plans, since mean plan does not make sense as a concept under the payoff framing presented to subjects, where each is given discrete strategies with accompanying payoff outcomes.

[8] The earnings function representing Table 1 is written as: $\pi_i\ (x_i, x_{-i}) = (30 - x_i) + Return \times (x_i + x_{-i} \times (N-1)) \div N$. x is a mapping from the choice of Plan to numerical correspondents. x_i is the numerical correspondent for each player i given her choice of plan, such that $x_i(Plan) = 30, 20, 10$ and 0 when $Plan = 1, 2, 3$ and 4, respectively. x_{-i} is the numerical correspondent for each player i given the most common choice of plan of all her direct neighbours. $x_{-i}(Modal\ Plan) = 30, 20, 10$ and 0 when $Modal\ Plan = 1, 2, 3$ and 4, respectively. The $Modal\ Plan$ is the most common plan chosen by player i's neighbours, however, if there is more than one plan chosen by neighbours most commonly, then the lowest one becomes the $Modal\ Plan$. Therefore, each subject's payoff depends on her choice of plan and the modal choice of plan of all her direct neighbours, in terms of the numerical correspondents. The $Return$ differs according to network size (N), therefore the marginal per capita return, which is defined by $Return/N$, decreases in network size. The table shown below summarizes the relationship between network size, $Return$ and marginal per capita return.

Network size (N)	1	2	3	4	5	6	7	8	9	10	11	12
$Return$	1.000	1.600	2.000	2.500	3.100	3.700	3.100	2.500	2.000	1.600	1.400	1.300
Marginal per capita return ($Return/N$)	1.000	0.800	0.667	0.625	0.620	0.617	0.443	0.313	0.222	0.160	0.127	0.108

[9] If the neighbours' modal plan is Plan 4, then the optimal number of neighbours is 0.

3.3 Treatment variations

There are two treatment variations in this experiment, *Integration* and *Mobility*. We employ a 2×2 treatment design, as summarized in Table 2.

Table 2: Treatments.

	Low Mobility (a 30% chance of being allowed to update a relationship)	High Mobility (an 80% chance of being allowed to update a relationship)
Low Integration (two starting neighbours from the in-group)	I30	I80
High Integration (two starting neighbours from the out-group)	O30	O80

Integration – This variation relates to the first round of the game. In Step 1 of Round 1, we fixed relationships for each pair.[10] In total, among the 12 participants, the number of pairs is 66. Among all 66 pairs, 12 pairs were fixed as neighbours, and the others were not. In the case of *Low Integration*, the 12 pairs of neighbours were determined such that every participant had two neighbours, and both neighbours had the same social identity as that participant. This resembles a society in which there is complete segregation between social identity groups. In the case of *High Integration*, the 12 pairs of neighbours were determined such that every participant had two neighbours, but both neighbours had the opposite social identity to the participant. This resembles a society where there is a high level of mixing between different groups. In both cases, each participant only knew her neighbours but did not know the neighbours of others, although the instructions told her that each participant was assigned two in-group or out-group neighbours.

Mobility – We varied the probability that each pair of participants was allowed to redefine their relationships in a given round. For each pair of participants, in each round the computer programme randomly, and independently, determined whether this pair's relationship was fixed or changeable. If it was fixed in the current round, the two participants automatically kept the same relationship as in the previous round, without either being allowed to initiate a link (if they were not previously neighbours) or break the connection (if they were). If the relationship was changeable, then both participants would be allowed to make such decisions, and their relationship could be redefined following the rules outlined in

[10] Therefore, participants were not allowed to adjust their relationships in Step 1 of Round 1. They started to determine their neighbours at Round 2, and onwards until Round 25.

the previous sub-section. Under *Low Mobility*, each relationship had a 30% probability of being changeable in a given round. Under *High Mobility*, this probability was 80%. This resembles two different types of societies, one in which it is relatively difficult for people to initiate new connections or break existing ones and therefore network structures are relatively inflexible, and the other in which such connections can be altered with relative ease and therefore network structures have the potential to be quite fluid.

3.4 Implementation

Prior to data collection, the design was pre-registered (Chen, 2023). The experiment was conducted in the CeDEx China Lab (University of Nottingham Ningbo China, UNNC) and CCBEF Lab (Southwest University of Finance and Economics, SWUFE) between February and March 2023. We implemented a between-subject design; each treatment had eight sessions, half of them conducted at UNNC and the other half at SWUFE. Since we have four treatments, in total there were 32 sessions.[11] Each session only contained 12 participants, i.e., one "society". Therefore, we recruited 384 participants, as described in Table 3. The recruitment went through ORSEE (Greiner, 2015) at UNNC and YZLabCloud[12] at SWUFE. Most of the participants were undergraduate students, 70.31% were female and 75.25% had Business or Social Sciences backgrounds. 77.34% of them had participated in either the CeDEx or CCBEF lab at least once before this experiment. Almost all participants reported they did not know other subjects in the same session. Only a small fraction of participants reported having some exposure to Bach's music used in this experiment (therefore for most of the participants, these music excerpts are meaningless in deciding their social identities).

The language used in the experiment was Chinese. The experimental instructions are translated from Chinese in Appendix 1. The experimental software was zTree (Fischbacher, 2007). Each session lasted about 90 minutes. Participants were paid in RMB. The average payment was 90.82 RMB (11.65 Euro),

[11] There were two extra sessions. One was a formal session, but it was cancelled because the number of subjects who showed up was less than the required number (we needed 12 subjects for each session). The other was a pilot session, designed to test whether the experimental procedure and instructions work well in the lab with subjects; the data of this session was not included in the data analysis of this paper because the payment scheme of this session was different from the later formal sessions.

[12] YZLabCloud is an online participant pool management system. It is different from other common choices (like ORSEE) since it is embedded within a WeChat sub-application (also called "mini-programme"). Therefore, participant recruitment for a session can be completed via WeChat instead of sending bulk email messages. This is suitable for Chinese universities because email is not frequently used as a communication device between the university and the students. Their web (in Chinese): http://www.yzlabcloud.com/.

with a minimum of 19.4 RMB (2.49 Euro), and a maximum of 132.5 RMB (17 Euro).

Table 3: Background information about participants.

	SWUFE (N = 192)		UNNC (N = 192)		Total (N = 384)	
Mean age	19.98	SD = 1.162	20.54	SD = 1.468	20.26	SD = 1.352
Is female	143	74.48%	127	66.15%	270	70.31%
Is undergraduate student	190	98.96%	184	95.83%	374	97.40%
Areas of study						
Arts	6	3.13%	12	6.25%	18	4.69%
Business	138	71.88%	60	31.25%	198	51.56%
Engineering	0	0.00%	28	14.58%	28	7.29%
Science	21	10.94%	28	14.58%	49	12.76%
Social Sciences	27	14.06%	64	33.33%	91	23.70%
Has prior experiment experience	174	90.63%	123	64.06%	297	77.34%
Has ever listened to VG or WT	5	2.60%	10	5.21%	15	3.91%
Mean number of friends in the lab	0.15	SD = 0.387	0.19	SD = 0.497	0.16	SD = 0.445

In each session, we started by introducing general information about the experiment. Next, we distributed the instructions for the music task, read it aloud and then conducted the music task. Then, we distributed the instructions for the discussion task, read it aloud and then asked participants to perform it. When the experiment reached the stage of the congestible public goods game, we told participants that the whole task would contain 25 rounds, and each round would contain two steps. Then we distributed the instructions for Step 1 (network formation), read it aloud, and asked participants to complete a set of comprehension questions; and repeated the same procedures for Step 2 (contribution decisions and outcomes). After that, we introduced who their neighbours were in Round 1, and participants started to complete the 25 rounds of the game. Throughout the experiment, participants did not know the content of the next task until they were given new instructions.

Participants were told that they would be paid by real money via mobile transfer at the end of the session. They were informed that for each round they would be told their own and their neighbours' choices of plan, as well as their own resulting payoffs, immediately at the end of that round, and that at the end of the experiment only one round would be randomly selected for payment.

4. Hypotheses

Regarding the effect of initial network structure, the treatments with initially integrated societies will naturally have a higher average proportion of in-group network members in early round of the experiment. Our research interest is in whether, after many opportunities to reshape the networks, initially segregated and initially integrated societies will ultimately converge to the same levels of homophily. This would be likely to occur if the level of in-group bias in link formation and deletion decisions was unrelated to the initial level of societal integration. However, this would not be likely to occur if, as seems plausible, early interactions in the congestible public goods game shape the linking preferences of participants. We therefore hypothesise:

Hypothesis 1*: Networks will remain more homophilic, even in the last round of the experiment, in the Low Integration treatments compared to the High Integration treatments.*

In a more fluid society, there are more opportunities to interact with new people. This increases the absolute amount of contact with out-group members. A large literature in psychology and economics argues that intergroup contact can weaken social identity, and reduce prejudice and discrimination (e.g., Allport, 1954; Freddi et al., 2024). Increasing opportunities for contact with the out-group might therefore also lead to less homophilic networks (Smith et al., 2014).

Hypothesis 2*: Networks will exhibit higher levels of homophily in the Low Mobility treatments than in the High Mobility treatments.*

5. Experimental results

5.1 Linking behaviour

5.1.1 Overall linking decisions

Table 4 describes the linking decisions of the sample across all sessions and treatments. On average, a participant has 4.340 (SD = 1.269) neighbours. Recall that the payoff of each participant depends on the number of neighbours she has, and having five neighbours earns the most for (almost) every choice combination. This may explain why the average number of neighbours we observed in our experiment is close to five. On average, each participant sends 1.499 (SD = 2.026) requests to form neighbour relationships and makes 0.255 (SD = 0.530) decisions to delete neighbours. Therefore, in general, we observe a high level of activity in attempts to form neighbours, while cases of exclusion (being kicked out of a network) are comparatively rare.

Table 4: General descriptive statistics on network formation.

	(1)	(2)	(3)		(4)		(5)	
	Total number of subjects	Total number of observations	Average number of neighbours per subject per round		Average number of links initiated per subject per round		Average number of links deleted per subject per round	
I80	96	2304	4.337	(1.159)	2.080	(2.531)	0.294	(0.427)
I30	96	2304	4.241	(1.215)	0.996	(1.242)	0.193	(0.554)
O80	96	2304	4.403	(1.330)	1.959	(2.429)	0.243	(0.470)
O30	96	2304	4.380	(1.356)	0.963	(1.216)	0.171	(0.636)
Total	384	9216	4.340	(1.269)	1.499	(2.026)	0.225	(0.530)

Note: Standard deviations are in parentheses. Round 1 is excluded because neighbours and links were determined exogenously.

We now present the linking data across rounds. Figure 1 shows the changes in the average numbers per round of neighbours, initiated links and deleted links in sequences of four rounds for each treatment. These figures only contain Rounds 2 to 25 because in Round 1 participants are not asked to determine their neighbours.

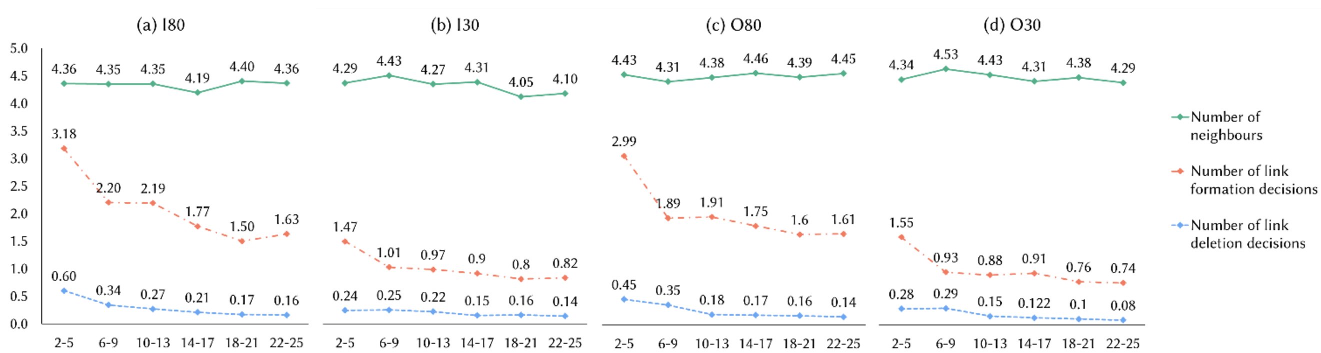


Figure 1: Average numbers per round of neighbours, link formation decisions and link deletion decisions in Rounds 2 to 5, 6 to 9, 10 to 13, 14 to 17, 18 to 21 and 22 to 25, by treatment.

The green solid line in each panel of Figure 1 shows that for all treatments, on average each participant has four to five neighbours, and this number almost does not fluctuate across rounds. The red dot-dash lines show that in all treatments the average number of initiated links decreases over the rounds, with the most rapid decline occurring early on. The blue dash line in each panel of Figure 1 shows a similar decline in the deletion of neighbours.

Overall, this indicates that in all treatments, networks become increasingly stable over time. In early rounds, participants are active in adjusting their network. They achieve this by initiating many links, but also deleting some of the links that they do not prefer. However, our data suggests that participants also try to control the number of neighbours in their network so that each is connected to close to five neighbours in each round. This leads to the general pattern that, once they have achieved a satisfactory network size, they add and delete links only to replace unsatisfactory neighbours, while maintaining the size of their network.

5.1.2 Determinants of linking decisions

We now consider the effects of group identity on link initiation. Recall that in Step 1 of the game, each participant can only decide to whom she wants to link given that the computer programme allows her to do so. To show to what extent participants are willing to invite others with the same social identity as themselves to be their neighbours, we calculate in each round each participant's linking rate towards the same-identity and the opposite-identity members who are currently non-neighbours and for whom the relationship is changeable. If the linking rate towards the same-identity members is larger than towards the opposite-identity members, then it means participants prefer the same-identity members to be their neighbours. The average linking rates of any subject in each round, averaging across Rounds 2 to 25, towards in-group candidates are 0.395 (SD = 0.453), 0.504 (SD = 0.484), 0.365 (SD = 0.439) and 0.495 (SD = 0.484) in I80, I30, O80 and O30, respectively. The corresponding average linking rates towards out-group candidates over Rounds 2 to 25 are 0.339 (SD = 0.415), 0.467 (SD = 0.480), 0.314 (SD = 0.403) and 0.441 (SD = 0.476) in I80, I30, O80 and O30, respectively.[13] In each treatment, the average linking rate towards in-group candidates is significantly higher than the linking rate towards out-group candidates (Wilcoxon signed-rank test, largest p = 0.017)[14].

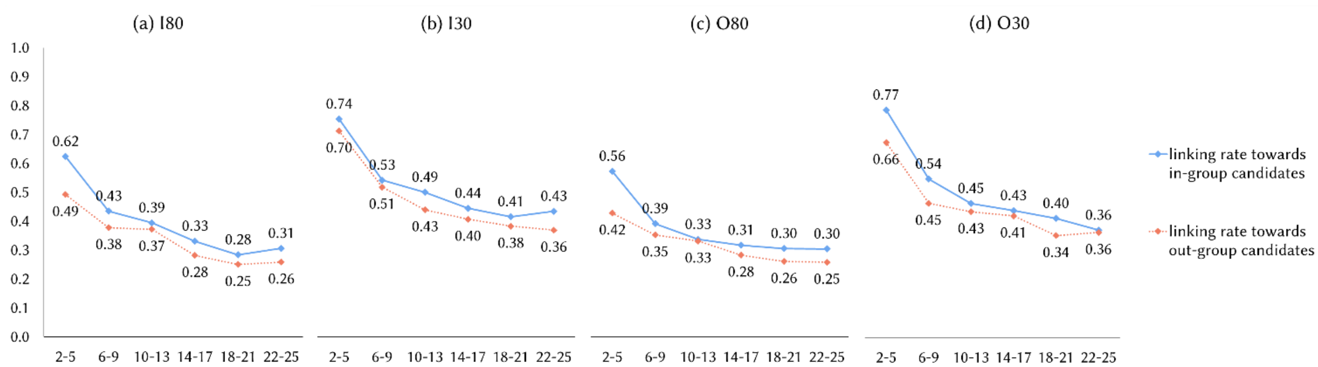


Figure 2: Linking rate towards in-group candidates and out-group candidates in Rounds 2 to 5,6 to 9, 10 to 13, 14 to 17, 18 to 21 and 22 to 25, by treatment.

Figure 2 shows the average linking rate towards in-group candidates and out-group candidates over time

---

[13] The linking rates towards in-group and out-group members are higher in treatments of low mobility. See analysis in Table A1 (Appendix 2).

[14] All significance tests in this paper are two-sided.

in each treatment. The two rates decrease over the rounds, but throughout participants generally remain more likely to initiate links to those of the same identity as themselves. Wilcoxon signed-rank tests show that in Rounds 2 to 5, the average linking rates are significantly higher toward same-identity candidates in all treatments (largest p = 0.001), while when the game arrives at Rounds 22-25, this significance only appears in O30 (p = 0.027; in other treatments, smallest p = 0.272). There is therefore some indication that this tendency for in-group favouritism shrinks over the course of the game.

Table 5: Link formation decisions.

Dependent variable: Link formation (conditional on receiving an opportunity)	Treatment			
	(1)	(2)	(3)	(4)
	I80	I30	O80	O30
Same Identity	0.717***	0.288*	0.567***	0.502***
	(0.129)	(0.154)	(0.103)	(0.122)
Number of Neighbours at r-1	-0.539***	-0.799***	-0.650***	-1.178***
	(0.147)	(0.237)	(0.137)	(0.205)
Round dummies				
r6-9	-0.315**	-0.526**	-0.155	-0.439
	(0.183)	(0.217)	(0.150)	(0.281)
r10-13	-0.300	-1.033***	-0.351**	-0.860***
	(0.278)	(0.211)	(0.150)	(0.227)
r14-17	-1.121***	-1.230***	-0.515***	-1.060***
	(0.391)	(0.196)	(0.137)	(0.298)
r18-21	-1.317***	-1.856***	-0.707**	-1.297***
	(0.342)	(0.367)	(0.284)	(0.329)
r21-25	-1.246***	-1.843***	-0.710***	-1.455***
	(0.290)	(0.289)	(0.228)	(0.275)
Constant	2.066***	4.227***	2.198***	5.546***
	(0.355)	(0.753)	(0.326)	(0.674)
Number of obs.	12,670	4,776	12,502	4,630
Number of clusters	8	8	8	8
Wald chi2	947.69	2133.88	415.26	6840.54
Prob > chi2	0.000	0.000	0.000	0.000
Log likelihood	-5641.005	-2442.752	-5334.700	-2175.824

Note: Random-effects logistic regressions. In the models, each set of decisions by one participant in one round represents a group of observations. Coefficients are reported, with standard errors in parentheses. Standard errors are clustered at the session level; each treatment has eight clusters. Data is only included from relationships in which link initiation is allowed. Dependent variable = 1 if a participant initiates a link to that neighbour. Dummy variable Same Identity = 1 if the proposed candidate has the same social identity as the decision maker. Round dummies are included. $p < 0.1$*, $p < 0.05$**, $p < 0.01$***.

To test the effect of social identity on link formation decisions, we report in Table 5 the results from a random-effects logit regression with the dependent variable being the link proposal decision (which equals one if the link is initiated). We conduct the regressions separately for each treatment. We cluster the robust standard errors at the session level and use observations from Rounds 2 to 25. The independent variables include (1) a dummy variable that equals one if the candidate has the same social identity as the decision maker and (2) the number of neighbours the decision maker has at the start of the round. These are the two pieces of information the decision-maker has on screen when making a link initiation decision. We also include round sequences as dummies to control for changes in linking behaviour throughout the experiment.

The main insight from Table 5 comes from the coefficient on Same Identity. It shows that given the opportunity of becoming neighbours, participants prefer those with the same social identity as themselves (largest p = 0.061 across the four treatments). Also, participants care about the number of neighbours in their network. Having more neighbours decreases her likelihood of initiating extra links (largest p = 0.001 across the four treatments). This is consistent with what we observed in Section 4.1.1 in which it appeared that participants try to control the number of neighbours to be around five. Furthermore, the coefficients of the round sequence dummies are negative, and are statistically significant, especially after Round 13, confirming the fact that participants become less active in initiating links in later rounds (largest p = 0.013 across the four treatments). In further models, presented in Table A3 (Appendix 3), we also interact the identity dummy with the Round dummies and the results from these models confirm that in-group favouritism decreases over the rounds.

Conclusion 1*: When participants decide who to link with as neighbours, they tend to invite others in the same social identity group as themselves. The likelihood of getting an invitation decreases in the number of neighbours the decision-maker has and decreases over time. These results apply to all treatments.*

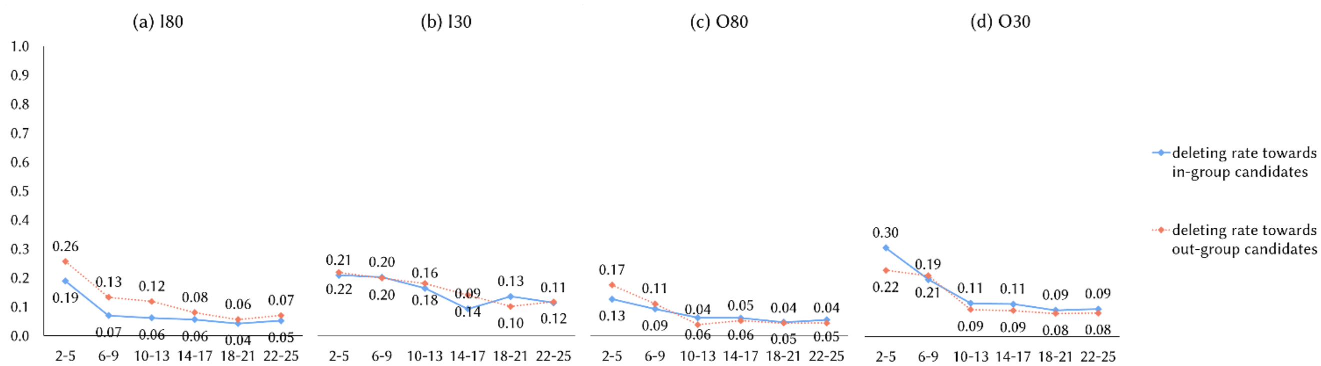


Figure 3: Deleting rate towards in-group candidates and out-group candidates in Rounds 2 to 5, 6 to 9, 10 to 13, 14 to 17, 18 to 21 and 22 to 25, by treatment.

We now investigate the determinants of decisions to delete a neighbour from one's network, especially the effects of social identity. Recall that only a random subset of current neighbours are available for deletion in a given round. We again calculate two rates – the deleting rate towards in-group neighbours and towards out-group neighbours, only including cases where changing the relationship is allowed in the current round. The deleting rates of each subject in each round towards in-group candidates are 0.078 (SD = 0.219), 0.152 (SD = 0.343), 0.070 (SD = 0.210) and 0.136 (SD = 0.329) in I80, I30, O80 and O30, respectively. The deleting rates of each subject in each round towards out-group candidates are 0.109 (SD = 0.273), 0.148 (SD = 0.334), 0.075 (SD = 0.220) and 0.128 (SD = 0.309) in I80, I30, O80 and O30, respectively.[15] Wilcoxon signed-rank tests show that the two rates are insignificantly different from each other except in I80 (p = 0.000). Figure 3 shows the average deleting rate towards in-group candidates and out-group candidates over time in each treatment. Wilcoxon signed-rank tests show that at Rounds 2 to 5, only participants in I80 are more likely to delete out-group than in-group neighbours (p = 0.065). In Rounds 22 to 25, the two rates of deletion are indifferent in all treatments (Wilcoxon signed-rank test, smallest p = 0.103). Therefore, there is evidence for discrimination in deletion decisions only in the I80 treatment, and even there only in earlier rounds.

To further test the effect of the same identity on the link deletion decision, we report in Table 6 the results from a random-effects logit regression with the dependent variable being the link deletion decision (which equals one if the participant deletes that neighbour from her network). The regression specifications and control variables used in the regressions shown in Table 6 are the same as in Table 5 for results of link formation decisions, except that we add a new independent variable Plan r-1. Plan r-1 represents the neighbour's choice of plan at the last round, which equals 1, 2, 3 or 4. When participants decide whether to delete a neighbour, this neighbour's last round choice, as well as her social identity, are displayed on screen.

The main insights from Table 6 are as follows. In the case of *High Mobility*, having the same social identity as the decision maker leads the neighbour to be kept in the social network. They are found to be significantly less likely than out-group neighbours to be deleted by decision-makers in models (1) and (3) (largest p = 0.099). In contrast, in the case of *Low Mobility*, social identity does not significantly affect the likelihood, as the coefficient on Same Identity is statistically insignificant in models (2) and (4) (smallest p = 0.259). Instead, a neighbour is more likely to be kicked out of the network when she chooses a more selfish plan (coefficient Plan r-1, largest p = 0.069 in models (2) and (4)), as well as when the decision maker has a relatively high number of neighbours (largest p = 0.039 in column (2) and column

[15] The deleting rates towards in-group and out-group members are higher in treatments of low mobility. See analysis in Table A2, (Appendix 2). This finding, along with that of footnote 13, demonstrates that when opportunities to reshape one's network are rarer, people take more advantage of available opportunities.

(4)).[16] The Round dummies are negative, demonstrating that the tendency to delete diminishes as the game is played repeatedly. In further models (presented in Table A4, Appendix 3), we also interact the identity dummies with the round dummies. The results show that only in treatment O80 does the effect of identity change significantly over time.[17]

Table 6: Link deletion decisions.

Dependent variable: Link deletion (conditional on receiving an opportunity)	Treatment			
	(1)	(2)	(3)	(4)
	I80	I30	O80	O30
Same Identity	-0.826***	-0.195	-0.372*	0.070
	(0.193)	(0.173)	(0.226)	(0.172)
Plan r-1	0.106	0.287**	-0.101	0.214*
	(0.066)	(0.137)	(0.162)	(0.118)
Number of Neighbours at r-1	0.073	0.220**	0.033	0.405***
	(0.054)	(0.106)	(0.072)	(0.072)
Round dummies				
r6-9	-0.745*	-0.077	-0.243	-0.459
	(0.423)	(0.151)	(0.388)	(0.347)
r10-13	-0.742**	-0.082	-1.034***	-1.266***
	(0.292)	(0.192)	(0.213)	(0.306)
r14-17	-1.152*	-0.550	-1.261***	-1.040***
	(0.620)	(0.348)	(0.316)	(0.212)
r18-21	-1.556***	-0.473	-1.247***	-1.404***
	(0.576)	(0.526)	(0.332)	(0.375)
r21-25	-1.479**	-0.691	-1.522***	-1.521***
	(0.634)	(0.632)	(0.452)	(0.424)
Constant	-2.199***	-3.400***	-2.607***	-3.964***
	(0.403)	(0.824)	(0.676)	(0.658)
Number of obs.	7,794	2,966	7,960	2,986
Number of clusters	8	8	8	8
Wald chi2	2458.52	515290.86	75750.32	4760.86
Prob > chi2	0.000	0.000	0.000	0.000
Log likelihood	-1800.752	-1128.178	-1566.809	-1002.785

Note: Random-effects logistic regressions. In the models, each set of decisions by one participant in one round represents a group of observations. Coefficients are reported, with standard errors in parentheses. Standard errors are clustered at the session level; each treatment has eight clusters. Data is only included from relationships in which link deletion is allowed. Dependent variable = 1 if a participant deletes a link with that neighbour. Dummy variable Same Identity = 1 if the neighbour has the same social identity as the decision maker. Plan r-1 represents the choice of Plan 1, 2, 3 or 4 (from the best to the worst) of the neighbour at r-1. Round dummies are included. $p < 0.1$*, $p < 0.05$**, $p < 0.01$***.

Conclusion 2*: When participants decide who to delete from their network, they tend to delete out-group neighbours in the case of High Mobility, and they tend to delete neighbours with selfish contribution*

[16] The effect of Same Identity on deletion likelihood is significantly more negative under high mobility and under low integration. The effect of Plan r-1 is smaller under both high integration and high mobility, but not quite significantly so. Details are presented in Table A5, Appendix 4.

[17] We also interact the identity dummy with the neighbour's previous round plan. However, this does not find evidence that deletions are more sensitive to the past decisions of in-group or out-group members. Details are presented in Table A6 (Appendix 5).

behaviour in the case of Low Mobility.

We now turn to discuss the effect of *Integration* and *Mobility* on the composition of network members. Figure 4 shows the changes in the average fraction of in-group neighbours over the rounds. The average fraction equals one for I80 and I30 in Round 1 because, in these two treatments, each participant is assigned two neighbours having the same social identity as herself. The average fraction equals zero for O80 and O30 at round one because the two neighbours assigned to each participant in these two treatments have the other social identity. The average fraction decreases in Rounds 2 to 5 in I80 and I30 and increases in O80 and O30 because, when participants are allowed to invite neighbours, they initiate links with both in-group and out-group players. After approximately Round 5, for all treatments, the average fraction shows some fluctuations but does not experience severe change.

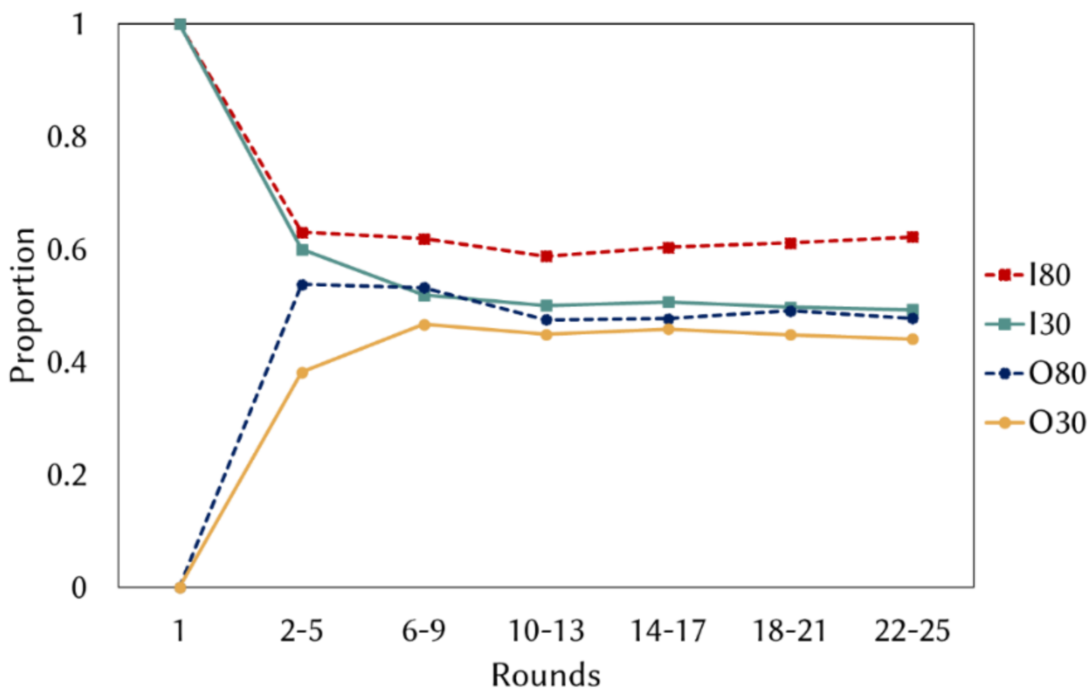


Figure 4: Average proportion of network members with the same identity in each round, by treatment.

Therefore, we calculate the average fraction of network members with the same identity in each treatment using Rounds 6 to 25. The result shows that on average, from Round 6 to Round 25, the fraction of same-social identity network members in treatment I30, I80, O30 and O80 are 0.503 (SD = 0.229), 0.609 (SD = 0.233), 0.453 (SD = 0.205) and 0.490 (SD = 0.213), respectively. We use an OLS regression to test for treatment effects, as shown in Table 7. The dependent variable is each subject's proportion of same-identity network members in each round. We include dummy variables for High Integration and High Mobility. The coefficient of High Mobility shows that in the two treatments I80 and O80, when each pair of relationships has a high probability of being allowed to be changed, there will be more of the same social identities gathered together. The coefficient of High Integration is negative, implying that, for treatments I80 and I30 in which the founding network members are of the same identity as the participant herself, there will be a stronger gathering of the same social identity members. The analysis shown in model (2) of Table 7 includes only Rounds 6-25; therefore, the effects of High Integration are not mechanically driven by the in-group relationships which are exogenously imposed at the start but persist after participants have had the opportunity to reorganise their networks. Moreover, these treatment effects exist in each of the round sequences we analyse separately in further regressions, as we show in models (3) to (5) of Table 7. Therefore, the influence of original segregation (the network initial condition) never disappears.

Table 7: Treatment effects on network homophily.

Dependent variable: Proportion of network members with the same social identity	Rounds				
	(1)	(2)	(3)	(4)	(5)
	Rounds 2-25	Rounds 6-25	Rounds 10-25	Rounds 14-25	Rounds 18-25
High Integration	-0.096***	-0.084***	-0.088***	-0.090***	-0.091***
	(0.021)	(0.025)	(0.026)	(0.028)	(0.029)
High Mobility	0.075***	0.071***	0.069**	0.073**	0.080**
	(0.021)	(0.025)	(0.026)	(0.028)	(0.029)
Round dummies					
r6-9	-0.003				
	(0.015)				
r10-13	-0.034*	-0.031***			
	(0.018)	(0.010)			
r14-17	-0.026	-0.023	0.009		
	(0.020)	(0.015)	(0.010)		
r18-21	-0.025	-0.022	0.009	0.0004	
	(0.019)	(0.015)	(0.011)	(0.009)	
r21-25	-0.029	-0.026*	0.005	-0.003	-0.004
	(0.019)	(0.014)	(0.011)	(0.011)	(0.009)
Constant	0.548***	0.541***	0.513**	0.520***	0.518
	(0.024)	(0.028)	(0.027)	(0.030)	(0.030)
# of Obs.	9,155	7,625	6,096	4,567	3,046
# of Clusters	32	32	32	32	32
F	9.20	7.27	8.24	9.25	9.52
R2 (adj.)	0.074	0.061	0.061	0.065	0.071
Prob.>F	0.000	0.000	0.000	0.000	0.000

Note: OLS regressions. Dependent variable is the proportion of network members with the same social identity. High Integration is a dummy variable that equals 1 for treatments O80 and O30 and equals 0 for treatments I80 and I30. High Mobility is a dummy variable that equals 1 for treatments I80 and O80 and equals 0 for treatments I30 and O30. Round 1 is excluded for all models. Round dummies are included; the omitted category in model (1) is Round 2-5, in model (2) is Round 6-9, in model (3) is Round 10-13, in model (4) is Round 14-17 and in model (5) is Round 18-21. Robust standard errors (in parentheses) are clustered at the session level. $p < 0.1^*$, $p < 0.05^{**}$, $p < 0.01^{***}$.

***Conclusion 3**: High Mobility and Low Integration work together to increase the average fraction of the same social identity network members throughout all rounds of play. Networks remain more homophilic in the Low Integration treatments compared to the High Integration treatments; Hypothesis 1 is supported. However, as networks are more homophilic in the High Mobility than the Low Mobility treatments, Hypothesis 2 is not supported.*

These treatment differences in homophily stem from the treatment differences documented above for the strength of in-group bias in deletion decisions. However, we cannot fully ascribe them to this single cause. Table A8 (Appendix 6) presents regression analysis showing that, with weak significance, existing neighbours were more likely to be randomly selected as eligible for deletion when they were in-group in O30, and when they were out-group in I80. These differences by neighbours' identity type in the frequency of available opportunities to delete them – which arose purely by luck – contributed mechanically to the low level of homophily in the former treatment and the high level in the latter. Thus, luck was a contributing factor in the magnitude – though not the direction – of treatment differences in homophily.

5.2 Contribution behaviour

5.2.1 Overall contribution decisions

Overall rounds, Plan 1 is chosen 59.03% of the time, while Plans 2, 3 and 4 are respectively chosen 11.05%, 10.47% and 19.45% of the time. Figure 5 shows the distribution of choices by treatment. It shows that in all treatments, Plan 1 is the most common choice. Figure A8 (Appendix 7) summarizes the distribution of plans by round in each treatment. It shows two features by round. First, in each treatment there is an increase in the choice of Plan 4 in the last round, similar to the so-called "end-game effect" in the standard public goods game (Andreoni, 1998). This suggests that the threat of eviction from neighbours' networks plays a disciplining role throughout the experiment until it is removed in the final round. Second, for the distribution of the plans, there are more subjects choosing higher plans in later rounds compared to earlier rounds, which means subjects contribute more to the public good game when time goes by. The regression results in Table 8 also confirm it. This time variation is consistent with other experimental studies about dynamic evolution of cooperation (see, e.g, Charness et al 2014; Riyanto and Teh, 2020).

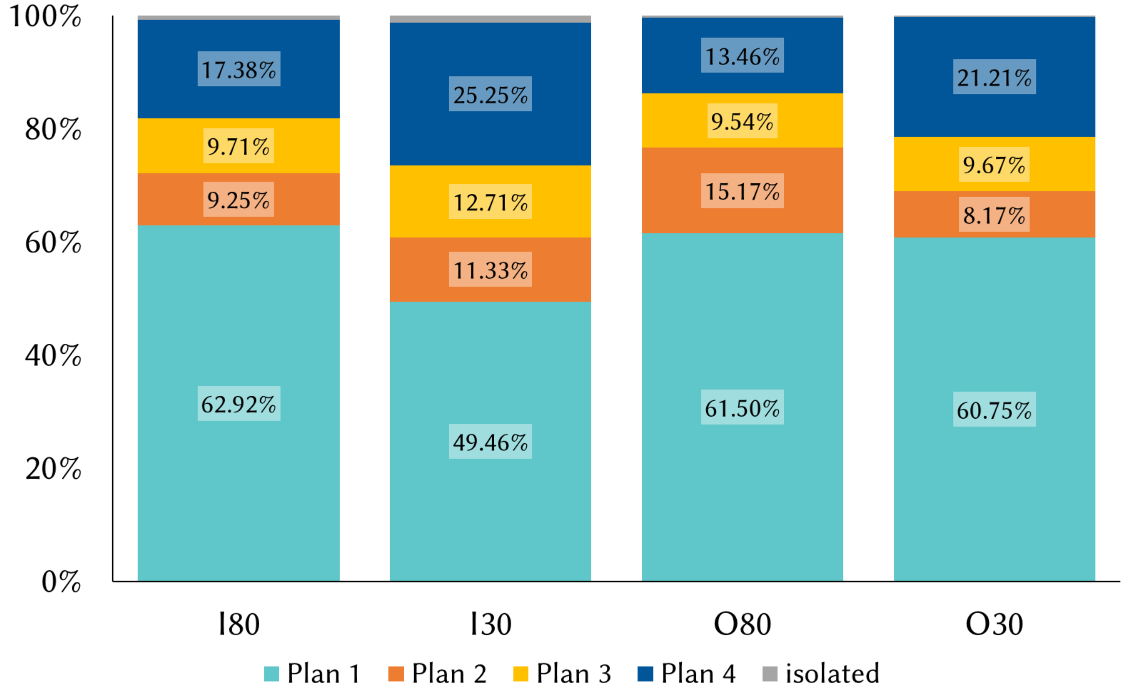


Figure 5: Distribution of Plans across treatments.
Note: Participants with zero neighbours in each round cannot make contribution decisions. These observations are marked as "isolated".

We also calculate the mean choice of plan in each round by converting the ordinal choice of Plan 1, Plan 2, Plan 3 or Plan 4 into the cardinal values 1, 2, 3 or 4, respectively. This represents a logical approach, since Plan 1, Plan 2, Plan 3 and Plan 4 are equivalent to contributing 100%, 66.6%, 33.3% and 0% of an initial endowment into the public good. Therefore, the distance between plans is constant.[18] Observations where subjects are not allowed to make contribution decisions due to having no neighbours are excluded. The mean choice of plan from Rounds 1 to 25 is 1.814 (SD = 0.024), 2.139 (SD = 0.026), 1.749 (SD = 0.022) and 1.913 (SD = 0.025) in treatment I80, I30, O80 and O30, respectively. Even the smallest mean (which appears in treatment I30), of the equivalent to 62.96% of the total initial endowment being contributed, is a comparatively high level of contribution compared to standard repeated public goods game, according to literature surveys (Chaudhuri, 2011; Kagel & Alvin, 1995).

[18] In the case that a mean plan is not an integer, this still has a clear interpretation – for instance, in I30 the mean plan is 2.13, which is equivalent to contributions being on average 62.96% of the initial endowment.

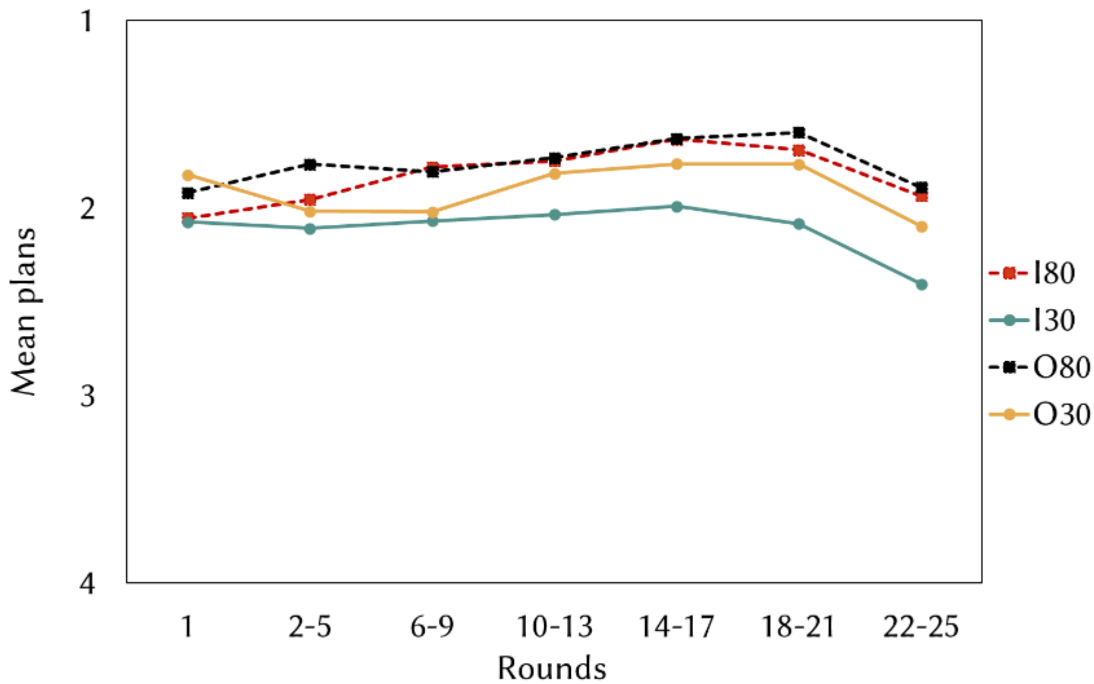


Figure 6: Mean contribution of plan by round.

5.2.2 Determinants of contribution decisions

We now turn to discuss the effect of *Integration* and *Mobility* on the level of contribution. We again use OLS regressions to test for treatment effects in Table 8. First, it shows that when all rounds (except Round 1) of play were included, both the coefficients of High Integration and High Mobility are statistically insignificant. In models (4) to (5) of Table 8, when the earlier rounds are excluded, it shows that the coefficient of High Mobility becomes statistically significant. Therefore, the treatment effect is statistically significant only in later rounds: when the game enters the last few rounds, subjects experiencing a higher rate of link adaption are selecting higher plans in interacting with their network members. Moreover, we find that subjects are more likely to choose a less selfish plan when the proportion of same-identity neighbours in her group becomes higher, and the treatment effects disappear when this is considered in the regression. When the data is separated by treatment, only in low integration treatments are subjects' choices of plan less selfish when they have more same-identity neighbours. Details are presented in Tables A10 and A11 (Appendix 8).

Conclusion 4*: The mean choice of plans is affected by the variations in Mobility. High Mobility leads subjects to select higher plans, however, it only affects subjects' contribution decisions in later rounds of* *the game. However, the initial in- and out-group matching does not bring significant changes on the choice of plans.*

Table 8: Treatment effects on choice of plans.

Dependent variable: Choice of plans of subjects	Rounds				
	(1)	(2)	(3)	(4)	(5)
	Rounds 2-25	Rounds 6-25	Rounds 10-25	Rounds 14-25	Rounds 18-25
High Integration	-0.128	-0.126	-0.154	-0.166	-0.191
	(0.158)	(0.156)	(0.159)	(0.162)	(0.171)
High Mobility	-0.251	-0.260	0.263	-0.289*	-0.310*
	(0.158)	(0.156)	(0.159)	(0.162)	(0.171)
Round dummies					
r6-9	-0.042				
	(0.039)				
r10-13	-0.128**	-0.087**			
	(0.054)	(0.041)			
r14-17	-0.207***	-0.165***	-0.079**		
	(0.070)	(0.059)	(0.037)		
r18-21	-0.176**	0.135*	-0.048	0.031	
	(0.077)	(0.070)	(0.048)	(0.030)	
r21-25	0.120	0.162**	0.249***	0.327***	0.297***
	(0.078)	(0.070)	(0.049)	(0.037)	(0.033)
Constant	2.150***	2.112***	2.040***	1.981***	2.035***
	(0.168)	(0.166)	(0.168)	(0.166)	(0.182)
# of Obs.	9,216	7,680	6,144	4,608	3,072
# of Clusters	32	32	32	32	32
F	17.85	20.54	24.21	29.80	36.15
R2 (adj.)	0.022	0.023	0.027	0.032	0.034
Prob.>F	0.000	0.000	0.000	0.000	0.000

Note: OLS regressions. Dependent variable is plan chosen by subjects. High Integration is a dummy variable such that it equals 1 for treatment O80 and O30 and it equals 0 for treatment I80 and I30. High Mobility is a dummy variable such that it equals 1 for treatments I80 and O80 and it equals 0 for treatments I30 and O30. Round 1 is excluded for all models. Round dummies are included; the omitted category in model (1) is Round 2-5, in model (2) is Round 6-9, in model (3) is Round 10-13, in model (4) is Round 14-17 and in model (5) is Round 18-21. Robust standard errors are clustered at the session level. p < 0.1*, p < 0.05**, p < 0.01***.

5.3 Payoff analysis

On average, each participant earns 81.547 (SD = 34.072), 71.799 (SD = 35.854), 84.712 (SD = 29.836) and 81.193 (SD = 34.639) in each round over the 25 rounds in I80, I30, O80 and O30. Figure 7 shows the average payoff of each treatment by round sequence. A regression confirms that there are not treatment differences, details are included in Table A12 (Appendix 9). There are two reasons to explain why payoffs do not differ across treatments, which can be understood with reference to the rules of determining payoffs that we introduced in Section 3.3. First, although in Section 4.2.2 we observed that subjects tend to choose higher plans in High Mobility treatments, it does not generally change the most common behaviour of neighbours of a subject – which is used to determine the payoff of each subject along with their own choices of plans (see Table A13, Appendix 10). Second, although in our setting payoffs also depend on network size, subjects in general had a very stable network size of around five. This means that in general their payoffs, in all treatments, are determined by the choice combinations from the Sub-table F (number of neighbours = 5) of Table 1 in Section 3.3.

***Conclusion 5**: The mean payoff was not affected by variations in Mobility or Integration.*

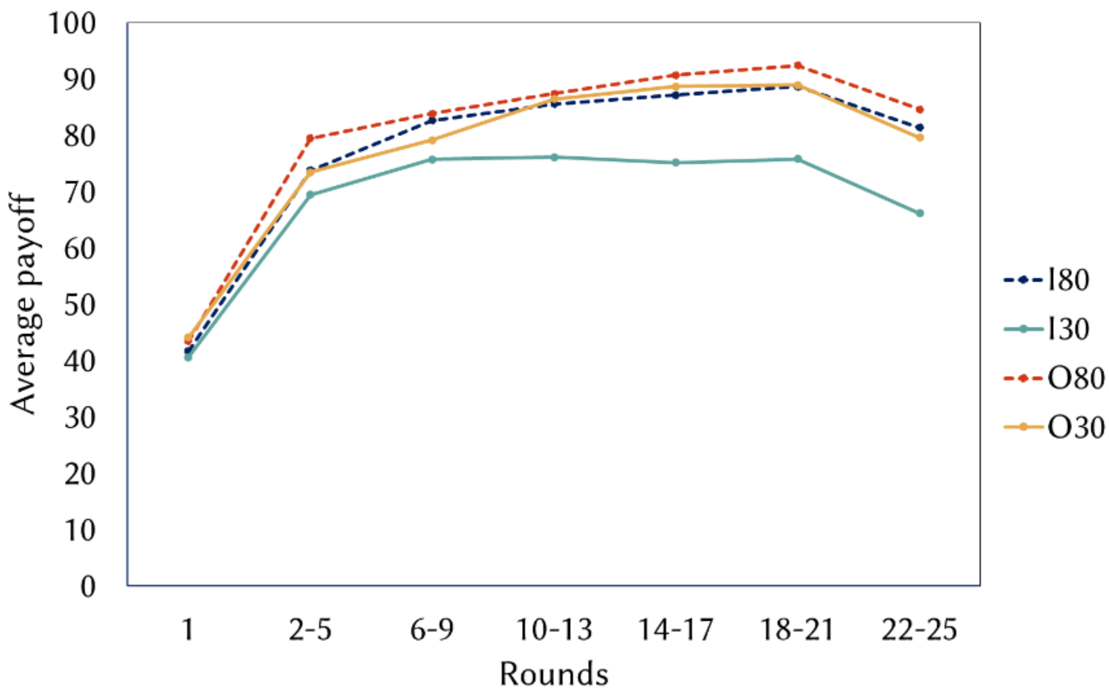


Figure 7: Changes in average payoff in RMB, by treatment.

6. Conclusion and discussion

In this paper, we experimentally investigate the role of social identity in network evolution in the context of a congestible public goods game, and how this varies according to the degree of societal mobility and the initial level of segregation. In our experiment, we identify in-group favouritism in the addition and deletion of network members, which in some cases persists across all rounds. Simply allowing repeated interactions does not guarantee the elimination of homophilic networks.

Our results also show that the degree of homophily in the evolving networks depends on the two treatment dimensions. First, network homophily is stronger with more segregated initial networks. On the one hand, this result highlights the importance of history and emphasises the difficulties that face policymakers attempting to integrate societies where segregation is initially entrenched: even after many repeated opportunities to reshuffle networks, which one might have hoped would render the arbitrary initial assignment irrelevant, the starting configuration still matters. On the other hand, from a more micro-based view on building a team or an organization (e.g., a starting cohort of employees), this result also implies the importance of initial deliberate interventions: during the induction period, assigning people to activities with out-group members should contribute to a more integrated organisation in the long run.

Second, network homophily is stronger as a result of higher mobility, which at first glance may seem a counterintuitive result. However, it provides us with a potential explanation why some huge cosmopolitan cities still end up with very high levels of social or residential segregation (Nilforoshan et al., 2023). Indeed, it is troubling that simply creating a more fluid society, allowing more opportunities for interactions which might weaken group identity, does not work in our setting and even seems to backfire, as individuals in fact use these opportunities to make their networks more homophilic. Meanwhile, in a less fluid society, people are more likely to delete free riders, perhaps because they fear being stuck with them for a long time in future if they do not take the immediate chance to remove them. Therefore, the strategies for reshaping networks applied in high and low mobility treatments seem different, such that subjects select neighbours based on in-group favouritism under high mobility but on contribution behaviour under low mobility. This backfiring effect of fluidity demonstrates the need for careful design of policies aimed at promoting societal integration.

Our results further show that, the patterns of in-group favouritism which determine network homophily appear in early rounds, when subjects are most active in rearranging their networks. In later rounds, players are less active in their adding and deleting behaviour, leading to relatively stable networks which retain the levels of homophily set by the crucial early interactions. Social identity is clearly important for first impressions, which have lasting consequences. However, it is possible that, once subjects are already connected and can observe each other's behaviour, willingness to cooperate becomes more salient over time than the identity that we assigned to them at the beginning of the experiment. Indeed, the inactivity in changing network structure in later rounds of the experiment might be due to people putting more weight on a new endogenized social identity based on interactions within the networks themselves. In Table A14 of Appendix 11, the regression results show supporting evidence that past contributions become increasingly important determinants of link deletion decisions over the rounds, especially in low integration treatments. Real-world examples could also reflect this point. For example, when recruiting new employees for a company, managers might care about whether the potential applicants share their social identity characteristics (via a submitted resume, interview and letters of recommendation), but once hired new employees may come to be regarded as members of the organizational in-group, and will be judged primarily on their work performance. Future studies could explore such a shift from exogenous to endogenous social identity, and its implications.

In this experiment, each subject has knowledge only of their own social network and does not have access to information about the networks of other agents. This setting replicates real-world conditions where individuals lack a comprehensive view of the global social network of their community. We believe the degree of network homophily could be increased if subjects were provided with information about the social networks of others. If there is a belief that individuals with a higher proportion of in-group members are more cooperative, or a conformist tendency to follow the homophilic behaviour observed among others, then the availability of public network information could push subjects to establish more connections with in-group rather than out-group members and finally exacerbate the clustering of in-group neighbours.

References

Ahn, T. K., Isaac, R. M., & Salmon, T. C. (2008). Endogenous group formation. *Journal of Public Economic Theory*, *10*(2), 171-194.

Ahn, T. K., Isaac, R. M., & Salmon, T. C. (2009). Coming and going: Experiments on endogenous group sizes for excludable public goods. *Journal of Public Economics*, *93*(1-2), 336-351.

Allport, G. W. (1954). *The nature of prejudice*. Cambridge/Reading, MA: AddisonWesley.

Allmis, P., & Merlino, L. P. (2023). Homophily and Specialization in Networks. *ArXiv preprint arXiv:2312.00457*.

Andreoni, J. (1988). Why free ride?: Strategies and learning in public goods experiments. *Journal of Public Economics*, *37*(3), 291-304.

Baccara, M., & Yariv, L. (2013). Homophily in peer groups. *American Economic Journal: Microeconomics*, *5*(3), 69-96.

Banerjee, A., Chandrasekhar, A. G., Duflo, E., & Jackson, M. O. (2013). The diffusion of microfinance. *Science, 341*(6144), 1236498.

Balliet, D., Wu, J., & De Dreu, C. K. (2014). Ingroup favoritism in cooperation: a metaanalysis. *Psychological Bulletin*, *140*(6), 1556.

Barnett, R. (2020). Leading with meaning: Why diversity, equity, and inclusion matters in US higher education. *Perspectives in Education*, *38*(2), 20-35.

Barr, A., Dekker, M., & Fafchamps, M. (2012). Bridging the gender divide: an experimental analysis of group formation in African villages. *World Development*, *40*(10), 2063-2077.

Bramoullé, Y., Currarini, S., Jackson, M. O., Pin, P., & Rogers, B. W. (2012). Homophily and long-run integration in social networks. *Journal of Economic Theory*, *147*(5), 1754-1786.

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics*, *14*, 47-83.

Charness, G., Cobo-Reyes, R., & Jiménez, N. (2014). Identities, selection, and contributions in a public-goods game. *Games and Economic Behavior*, *87,* 322-338.

Charness, G., Rigotti, L., & Rustichini, A. (2007). Individual behavior and group membership. *American Economic Review*, *97*(4), 1340-1352.

Charness, G., & Yang, C. L. (2014). Starting small toward voluntary formation of efficient large groups in public goods provision. *Journal of Economic Behavior & Organization*, *102*, 119-132.

Chen, Y. (2023). *How social identity affect network formation and public good provision?.* AEA RCT Registry. February 13. https://doi.org/10.1257/rct.10922-1.0

Chen, Y., & Li, S. X. (2009). Group identity and social preferences. *American Economic Review*, *99*(1), 431-457.

Coricelli, G., Fehr, D., & Fellner, G. (2004). Partner selection in public goods experiments. *Journal of Conflict Resolution*, *48*(3), 356-378.

Currarini, S., Matheson, J., & Vega-Redondo, F. (2016). A simple model of homophily in social networks. *European Economic Review*, *90*, 18-39.

Currarini, S., & Mengel, F. (2016). Identity, homophily and in-group bias. *European Economic Review*, 90, 40-55.

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

Davies, B. (2022). Gender sorting among economists: Evidence from the NBER. *Economics Letters*, *217*, 110640.

Ductor, L., & Prummer, A. (2024). Gender homophily, collaboration, and output. *Journal of Economic Behavior & Organization*, *221*, 477-492.

Ehrhart, K. M., & Keser, C. (1999). *Mobility and cooperation: On the run* (No. 99-69). Sonderforschungsbereich 504, Universität Mannheim & Sonderforschungsbereich 504, University of Mannheim.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171-178.

Friebel, G., Lalanne, M., Richter, B., Schwardmann, P., & Seabright, P. (2021). Gender differences in social interactions. *Journal of Economic Behavior & Organization*, *186*, 33-45.

Freddi, E., Potters, J., & Suetens, S. (2024). The effect of brief cooperative contact with ethnic minorities on discrimination. *Journal of Economic Behavior & Organization*, *222*, 64-76.

Gallo, E., Riyanto, Y. E., Roy, N., & Teh, T. H. (2022). Cooperation and punishment mechanisms in uncertain and dynamic social networks. *Games and Economic Behavior*, *134*, 75-103.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association*, *1*(1), 114-125.

Hargreaves Heap, S. P., & Zizzo, D. J. (2009). The value of groups. *American Economic Review*, *99*(1), 295-323.

Jackson, M. O. (2008). *Social and economic networks*. Princeton: Princeton university press.

Jackson, M. O. (2014). Networks in the understanding of economic behaviors. *Journal of Economic Perspectives*, *28*(4), 3-22.

Jackson, M. O., Nei, S. M., Snowberg, E., & Yariv, L. (2023). *The dynamics of networks and homophily* (No. w30815). National Bureau of Economic Research.

Kagel, J., & Alvin, R. (1995). Public goods: a survey of experimental research. *Handbook of Experimental Economics*.

Li, S. X. (2020). Group identity, ingroup favoritism, and discrimination. *Handbook of labor, human resources and population economics*, 1-28.

Mengel, F. (2020). Gender differences in networking. *The Economic Journal*, *130*(630), 1842-1873.

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

Nilforoshan, H., Looi, W., Pierson, E., Villanueva, B., Fishman, N., Chen, Y., ... & Leskovec, J. (2023). Human mobility networks reveal increased segregation in large cities. *Nature*, *624*(7992), 586-592.

Rand, D. G., Arbesman, S., & Christakis, N. A. (2011). Dynamic social networks promote cooperation in experiments with humans. *Proceedings of the National Academy of Sciences*, *108*(48), 19193-19198.

Riyanto, Y. E., & Teh, T. H. (2020). Highly flexible neighborhood promotes efficient coordination: Experimental evidence. *European Economic Review*, *129*, 103521.

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

Tan, J. H., & Bolle, F. (2007). Team competition and the public goods game. *Economics Letters*, *96*(1), 133-139.

Schelling, T. C. (1969). Models of segregation. *American Economic Review*, *59*(2), 488-493.

Small, M. L., & Pager, D. (2020). Sociological perspectives on racial discrimination. *Journal of Economic Perspectives*, *34*(2), 49-67.

Smith, J. A., McPherson, M., & Smith-Lovin, L. (2014). Social distance in the United States: Sex, race, religion, age, and education homophily among confidants, 1985 to 2004. *American Sociological Review*, *79*(3), 432-456.

van Leeuwen, B., Ramalingam, A., Rojo Arjona, D., & Schram, A. (2019). Centrality and cooperation in networks. *Experimental Economics*, *22*, 178-196.

Zeltzer, D. (2020). Gender homophily in referral networks: Consequences for the medicare physician earnings gap. *American Economic Journal: Applied Economics*, *12*(2), 169-197.

Appendix

A1. Experimental instructions

The experimental instructions are presented here, which have been translated into English. Treatment variations are highlighted in green and contents related to different laboratories (CeDEx and CCBEF) are highlighted in blue. The instructions need colour printing.

General introduction

Welcome to [In UNNC: CeDEx] [In SWUFE: CCBEF]! You are participating in an economics experiment. This experiment studies how individuals make decisions and the results will be used for scientific research. You can earn real money by completing tasks during the experiment. It is therefore very important that you read the following instructions carefully.

Rules of Conduct

We kindly ask you to follow the rules of conduct in the lab.

Before the experiment gets started, please switch off your cell phone and put it into the basket on the upper right of your screen. Throughout the whole experiment, please do not use cell phones and all other digital devices, please do not communicate with other participants except when instructed to, please do not save any screenshots that are relevant to the content of this experiment, and please do not focus on any other activities that are irrelevant with the experiment. Participants intentionally violating the rules will result in an immediate exclusion from the experiment and will not be paid.

If you have any questions during the experiment, please raise your hand. The experimenter will approach you to answer your questions.

Anonymity

This experiment will be conducted on PCs. All necessary communications among participants will take place via those networked computers.

All participants' personal identities will remain anonymous. It means no other participant learns about the identity of a participant that made a certain decision. It also means the payment procedure in this experiment will remain anonymous, as nobody will know other participants' final payment. It also means all personal identities will not be disclosed when we use the data of this experiment in research analysis.

Structure of the Experiment, and Payment

This experiment consists of three parts: Part One, Part Two and Part Three. You first receive the instructions for Part One and complete the required tasks. You next receive the instructions for Part Two and complete the required tasks. You then receive the instructions for Part Three and complete the required tasks. In the end, you need to complete a questionnaire.

Your earnings in the experiment partly depend on the choices made by yourself, partly on the choices made by other participants of today's session, and partly on a bit of chance. You will not know your earnings until the end of the experiment.

Your participation fee is 10 RMB. Your final payment is a sum of the participation fee and your earnings in the experiment. You will be paid at the end of the experiment.

If you have any questions, please raise your hand.

Part One

This part contains 5 questions. You need to complete each question independently.

In each question, you will hear 2 classical music excerpts. One excerpt comes from *Variations Goldberg* (henceforth, *VG*) and the other comes from *The Well-Tempered Clavier, Book II* (henceforth, *WT*), both of them are J.S Bach's compositions. Both excerpts are performed by the same pianist. You are asked to decide which excerpt you prefer.

In each question, the 2 excerpts are named Excerpt I and Excerpt II. This means, in listening to the music:

You will not be told which composition an excerpt comes from；

The order of appearance of the excerpt from *VG* and the excerpt from *WT* is random.

Each excerpt lasts for about 25 seconds, and each excerpt is played only 1 time. The experimenters will remind you before they start to play each except, and after it ends.

In each question, after the two excerpts are played, you are given 9 points. You need to allocate the 9 points to Excerpt I and Excerpt II according to your own preference over them:

- If you think Excerpt I is better than Excerpt II, you allocate more points to Excerpt I;
- If you think Excerpt II is better than Excerpt I, you allocate more points to Excerpt II.

When you are asked to make decisions, the decision page looks like as below:

(Example 1: A decision page, Part 1.)

There are 10 point-allocation options on the screen. Please choose one of them based on your preference for the two excerpts. You have 20 seconds to complete your decision. The countdown is displayed in the upper right corner of the screen.

After answering all 5 questions, the computer programme will calculate the total points you allocate to the 5 excerpts of *VG*, and the total points you allocate to the 5 excerpts of *WT*, respectively. In other words, since there are 5 questions which each is worth 9 points, you are given a total number of 45 points; the computer programme will detect how you allocate the 45 points between *VG* and *WT*.

Next, the computer programme will order all 12 participants of today's session by their total points toward *VG* from the highest to the lowest, and divide all participants into 2 music groups according to this order. Participants who are ordered between 1 and 6 are assigned to the **Variations group**, and those who are ordered between 7 and 12 are assigned to the **Well-Tempered group**.

For illustration purposes, the below table shows an example of assigning groups.

Column number / row number	Column 1	Column 2	Column 3	Column 4	Column 5
	participation code	total *VG*'s Points	total *WT*'s points	ordering according to the total *VG*'s points, from the highest to the lowest	group classification
Row 1	ckfc8im	45	0	1	**Variations group**
Row 2	knf2hi4	42	3	2	
Row 3	sh7ds8f	40	5	3	
Row 4	dif9dus	35	10	4	
Row 5	s6as87e	33	12	5	
Row 6	oaha6sa	30	15	6	
Row 7	efa7e9a	25	20	7	**Well-Tempered group**
Row 8	w89lda7	19	26	8	
Row 9	feaf75e	16	29	9	
Row 10	i9qwe5e	15	30	10	
Row 11	ldj8sm7	10	35	11	
Row 12	for4e3w	5	40	12	

(Example 2: A group allocation, Part 1.)

The computer programme will tell you whether you are assigned to the **Variations group** or the **Well-Tempered group** at the end of this task immediately.

If you are assigned to the **Variations group**, your group information **【You belong to the Variations group】** will appear on the upper right corner of the screen.

If you are assigned to the **Well-Tempered group**, your group information **【You belong to the Well-Tempered group】** will appear on the upper right corner of the screen.

If you have any questions, please raise your hand.

Part Two

This part contains 5 questions. You need to complete each question with the 5 group members of your music group. Each group member gets the same questions.

To find the answers, you and your group members can discuss. The discussion page looks like below:

(Example 3: A discussion page – take the *Well-Tempered* group as an example, Part 2.)

Please input your discussion in the purple dialogue box of the discussion page and then press "Enter" on the keyboard to send it to all group members. Your group discussion lasts for 600 seconds. The countdown is displayed in the upper right corner of the screen. Note that when you are talking with your group members, you will be assigned a temporary ID 【Group name + anonymity + a random number】.

When group discussion is completed, you need to submit the answer to each question on the next decision page. You have 90 seconds. The countdown is displayed in the upper right corner of the screen. When all participants finish their submission:

The 6 group members who are assigned to the *Variations* group will randomly get one of the below unique IDs of the *Variations* group:

【VG1】【VG2】【VG3】【VG4】【VG5】【VG6】

The 6 group members who are assigned to the *Well-Tempered* group will randomly get one of the below unique IDs of the *Well-Tempered* group:

【WT1】【WT2】【WT3】【WT4】【WT5】【WT6】

You will make use of this unique ID to complete tasks in the remainder of this session.

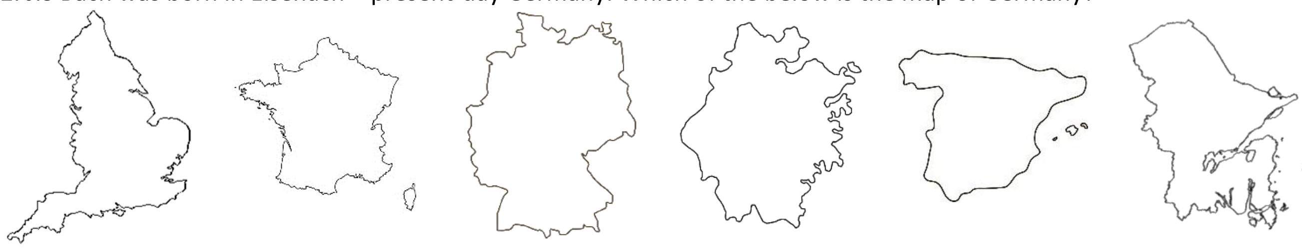
If you have any questions, please raise your hand.

Below are the 5 questions, and they are shown on the right-hand side of your discussion screen.

1. J.S Bach is the composer of all music excerpts in Part One. Which of the below is(are) portrait(s) of him?

A B C D E F

2. J.S Bach was born in Eisenach – present-day Germany. Which of the below is the map of Germany?



A B C D E F

3. Piano is used to perform all music excerpts in Part One. Which of the below is(are) a piano(s)?

A B C D E F

4. What is the sum of the last three digits of all 6 group members' cell phone numbers?

5. How many of all 6 group members went to the [In UNNC: No.1 Canteen] [In Chengdu: Wugu Canteen] for breakfast this morning?

<u>If you have any questions, please raise your hand.</u>

Part Three

You and the other 11 participants are asked to have 25 rounds of interactions. The rounds are identical. Each round takes 2 steps:

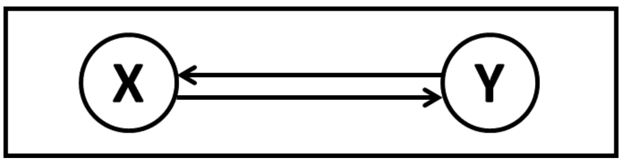
Step 1: You need to determine who are your neighbours (the other 11 participants **are not necessarily** all your neighbours);

Step 2: You and your neighbours can generate earnings by choosing plans.

We now introduce how to determine your neighbours in Step 1, and how to choose plans in Step 2.

Step 1

A pair of neighbours refers to 2 participants who are linked to each other. The example below helps illustrate this. Participant X sends a link to Participant Y, and Participant Y sends a link to Participant X. Therefore, Participant X becomes Participant Y's neighbour, and Participant Y becomes Participant X's neighbour.



(Example 4: A pair of neighbours.)

In this step, the computer programme will give you opportunities to add new neighbours, and delete current neighbours. The decision page looks like below.

(Example 5: A decision page, Part 3.)

The decision page consists of Windows 1, 2, 3 and 4. Now we introduce each of them.

<u>Window 1 – Upper left of the decision page</u>

(Example 6: a Window 1, Part 3)

In Window 1, you need to decide whether you want to delete links with **some of** your current neighbours.

In each round, for the participants who have been your neighbours in the previous round, each of them will have a **30%** chance [In the other treatment variation: a **80%** chance] to appear in this window.

In each row, the first column shows the ID of the neighbour, and the second column shows that neighbour's Plan in Step 2 of the previous round (we will explain what are the Plans very soon). In the third column, you need to choose either "Cancel" or "Not Cancel" the link to that neighbour. "Cancelling" a link means you do not want that participant to be your neighbour in this round. "Not Cancelling" a link means you want to maintain the neighbourship with that participant in this round.

When you are deciding whether to maintain the relationship with a certain neighbour, that neighbour is considering the same question as you, i.e., they need to decide to either "Cancel" or "Not Cancel" the link to you. If **either of** you chooses to "Cancel" the link, you both will not be neighbours with each other in this round.

The result of your mutual decisions will be shown in the fourth column.

- "Link is cancelled" means **at least one of** you chooses to "Cancel" the link. Therefore, you are not neighbours with each other in this round.
- "Link is not cancelled" means you **both** choose to "Not Cancel" the link. Therefore, you will keep the neighbourship in this round.

Window 2 – Upper right of the decision page

是否建立邻居关系:	你想发送链接与否?	你俩共同决定的结果
Whether to build links with the below participants	*Do you want to send the link or not?*	*Result of your two's mutual decisions*
【平3】【WT3】	发送 ○ *Send* ○ 不发送 *Not Send*	
【平6】【WT6】	发送 ○ *Send* ○ 不发送 *Not Send*	
【变1】【VG1】	发送 ○ *Send* ○ 不发送 *Not Send*	
【变2】【VG2】	发送 ○ *Send* ○ 不发送 *Not Send*	

(Example 7: A Window 2, Part 3.)

In this table, you need to decide whether you want to form links with **some of** the participants who are not your neighbours.

In each round, for the participants who have not been your neighbours, each of them will have a **30%** chance [In the other treatment variation: an **80%** chance] to appear in this window.

In each row, the first column shows the ID of the participant, and you need to choose to either "Send" or "Not Send" the link to this participant in the second column. "Sending" a link means you would like this participant to be your neighbour in this round. "Not Send" a link means you do not want this participant to be your neighbour in this round.

When you are deciding whether to develop a relationship with a certain participant, that participant is considering the same question as you, i.e.,they need to decide to either "Send" or "Not Send" the link to you. Only if **both of you** choose to "Send" the link, you both can be neighbours with each other in this round.

The result of your mutual decisions will appear in the third column.

- "Link is built" means you **both** chose to "Send" the link. Therefore, you are going to be neighbours with each other in this round.
- "Link is not built" means **at least one of** you chose to "Not Send" the link. Therefore, you both fail to develop the neighbourship in this round.

In Window 1 and Window 2, you need to make link decisions appear on every row, you cannot leave blanks. However, it might be the case that in some of the rounds, nothing will appear in one of the windows (or both windows). This means the computer programme does not give you opportunities to add and/or delete neighbours. If it happens, you do not have to make decisions in Step 1.

Window 3 and Window 4 – Bottom of the decision page

Once you have made decisions in Window 1 and Window 2, please click “OK” in Window 4 to submit your replies.

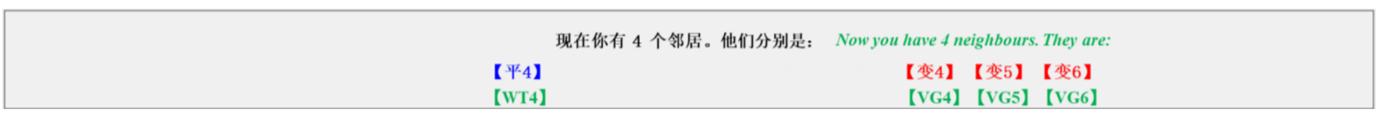
(Example 8: A Window 4, Part 3.)

All participants make decisions at the same time in this step. This means, you do not have a chance to discuss with other participants.

You have 40 seconds to make decisions. The countdown is displayed in the upper right corner of the screen.

After all of the participants submit their decisions, Window 3 will update:

- how many neighbours you obtain, and
- their IDs.



(Example 9: An example of Window 3, Part 3)

<u>If you have any questions, please raise your hand.</u>

To confirm your understanding of the instructions, please complete the comprehension questions. The decisions and payoffs used for the questions below are simply for illustrative purposes. Your answers will not affect your earnings in the following tasks.

Now, suppose your unique group ID is **【WT1】**, and please answer the below questions.

1. Suppose in a round, the computer programme asks you whether you want to keep a neighbourship with [VG1] in Window 1. Your decision is "Not Cancel". When you submit, the result is shown below:

Whether to cancel links with the below neighbours	This neighbour's choice over plans last round	Do you want to cancel the link or not?	Result of your two's mutual decisions
【VG1】	Plan 1	Cancel ○ ◉ Not Cancel	Link is not cancelled

1.1 Will [VG1] still be your neighbour in this round? A. Yes. ✓ B. No. ✗
1.2 Will you still be [VG1]'s neighbour? A. Yes. ✓ B. No. ✗
1.3 What is [VG1]'s decision in this round? A. [VG1] chooses "Cancel". ✗ B. [VG1] chooses "Not Cancel". ✓

2. Suppose in a round, the computer programme asks you whether you want to keep a neighbourship with [VG1] in Window 1. Your decision is "Not Cancel". When you submit, the result is shown below:

Whether to cancel links with the below neighbours	This neighbour's choice over plans last round	Do you want to cancel the link or not?	Result of your two's mutual decisions
【VG1】	Plan 1	Cancel ○ ◉ Not Cancel	Link is cancelled

2.1 Will [VG1] still be your neighbour in this round? A. Yes. ✗ B. No. ✓
2.2 Will you still be [VG1]'s neighbour? A. Yes. ✗ B. No. ✓
2.3 What is [VG1]'s decision in this round? A. [VG1] chooses "Cancel". ✓ B. [VG1] chooses "Not Cancel". ✗

3. Suppose in a round, the computer programme asks you whether you want to form a neighbourship with [VG1] in Window 2. Your decision is "Send". When you submit, the result is shown below:

Whether to build links with the below participants	Do you want to send the link or not?	Result of your two's mutual decisions
【VG1】	Send ◉ ○ Not Send	Link is built

3.1 Will [VG1] be your neighbour in this round? A. Yes. ✓ B. No. ✗
3.2 Will you be [VG1]'s neighbour? A. Yes. ✓ B. No. ✗
3.3 What is [VG1]'s decision in this round? A. [VG1] chooses "Send". ✓ B. [VG1] chooses "Not Send". ✗

4. Suppose in a round, the computer programme asks you whether you want to form a neighbourship with [VG1] in Window 2. Your decision is "Send". When you submit, the result is shown below:

Whether to build links with the below participants	Do you want to send the link or not?	Result of your two's mutual decisions
【VG1】	Send ◉ ○ Not Send	Link is not built

4.1 Will [VG1] be your neighbour in this round? A. Yes. ✗ B. No. ✓
4.2 Will you be [VG1]'s neighbour? A. Yes. ✗ B. No. ✓
4.3 What is [VG1]'s decision in this round? A. [VG1] chooses "Send". ✗ B. [VG1] chooses "Not Send". ✓

5. If in a certain round of Step 1, your decision screen page looks like the below:

5.1 How many neighbours do you have in this round?

A. 3 ✗ B. 4 ✓ C. 7 ✗

5.2 Which (one or more) of the following neighbours offered by the computer programme can be considered by you to be deleted in this round?

A. [WT4] [VG4] [VG6] ✗ B. [WT4] [VG4] [VG5] ✓ C. [VG1] [VG4] [VG5] ✗

5.3 Which (one or more) of the following participants offered by the computer programme can be considered by you to be your neighbour(s) in this round?

A. [WT3] [WT4] [VG1] [VG2] ✗ B. [WT3] [WT4] [VG1] [VG5] ✗ C. [WT3] [WT6] [VG1] [VG2] ✓

5.4 Who is your neighbour and the computer programme does not offer you a chance to consider whether to cancel her or not in this round?

A. [WT4] ✗ B. [VG4] ✗ C. [VG6] ✓

5.5 Who is not your neighbour and the computer programme does not offer you a chance to consider whether to build a link with her or not in this round?

A. [WT3] ✗ B. [WT5] ✓ C. [VG1] ✓

<u>If you have any questions, please raise your hand.</u>

<u>Correct answers and explanations have been shown on the right-hand side of the screen. Please check your responses. If you have any questions, please raise your hand.</u>

Question 1：

<u>1.1 A 1.2 A</u> Column 4 “Result of your two’s mutual decisions” shows “Link is not cancelled”, therefore, [VG1] is still your neighbour and you are still [VG1]’s neighbour in this round.

<u>1.3 B</u> You chose “Not Cancel” in Column 3, and in Column 4 “Result of your two’s mutual decisions” shows “Link is not cancelled”, therefore, your neighbour [VG1] chose “Not Cancel” in Column 3 as well. Only if you both choose “Not Cancel”, the information in Column 4 will be “Link is not cancelled”.

Question 2：

<u>2.1 B 2.2 B</u> Column 4 “Result of your two’s mutual decisions” shows “Link is cancelled”, therefore, [VG1] is not your neighbour and you are not [VG1]’s neighbour in this round.

<u>2.3 A</u> You chose “Not Cancel” in Column 3, and in Column 4 “Result of your two’s mutual decisions” shows “Link is cancelled”, therefore, your neighbour [VG1] chose “Cancel” in Column 3. As long as one of you chooses “Cancel”, the information in Column 4 will be “Link is cancelled”.

Question 3:

<u>3.1 A 3.2 A</u> Column 3 “Result of your two’s mutual decisions” shows “Link is built”, therefore, [VG1] becomes your neighbour and you become [VG1]’s neighbour in this round.

<u>3.3 A</u> You chose “Send” in Column 2, and in Column 3 “Result of your two’s mutual decisions” shows “Link is built”, therefore, your neighbour [VG1] chose “Send” in Column 2 as well. Only if you both choose “Send”, the information in Column 3 will be “Link is built”.

Question 4:

<u>4.1 B 4.2 B</u> Column 3 “Result of your two’s mutual decisions” shows “Link is not built”, therefore, [VG1] does not become your neighbour and you do not become [VG1]’s neighbour in this round.

<u>4.3 A</u> You chose “Send” in Column 2, and in Column 3 “Result of your two’s mutual decisions” shows “Link is not built”, therefore, your neighbour [VG1] chose “Not Send” in Column 2. As long as one of you chooses “Not Send”, the information in Column 4 will be “Link is not formed”.

Question 5:

<u>5.1 B</u> Check Window 3 — “Now you have 4 neighbours……”, therefore, now the number of your neighbours is 4.

<u>5.2 B</u> Check Window 1 — The computer programme asks you whether you want to delete neighbour [WT4], [VG4] and [VG5].

<u>5.3 C</u> Check Window 2 — The computer programme asks you whether you want to build links with [WT3], [WT6], [VG1] and [VG2].

<u>5.4 C</u> Check Window 3 first — “…… they are [WT4], [VG4], [VG5], [VG6]”; and check Window 1 next — the neighbours the computer programme asks you whether you want to delete are [WT4], [VG4] and [VG5]; therefore, the computer programme does not give you chance to delete [VG6].

<u>5.5 B</u> Check Window 3 first — “…… they are [WT4], [VG4], [VG5], [VG6]”, therefore, [WT5] is not your neighbour; and check Window 2 next — the computer programme asks you whether you want to build links with [WT3], [WT6], [VG1] and [VG2], therefore, [WT5] is not your neighbour and you do not have chance to build a link with her.

Step 2

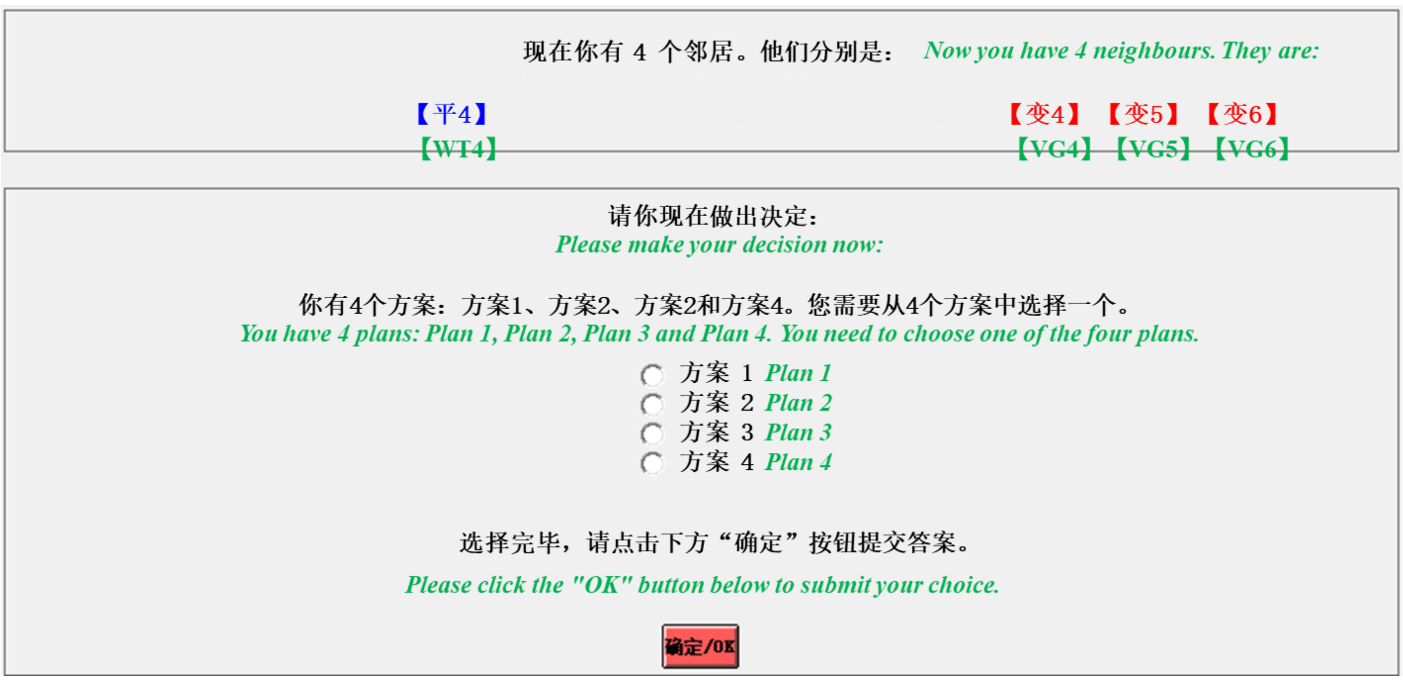
Next you can generate earnings with your neighbours as updated in Window 3. The decision is described below.

You have four plans: Plan 1, Plan 2, Plan 3 and Plan 4. You need to choose one of the four plans.

Like you, each of your neighbours is given these four plans, and needs to choose one among them.

All participants make decisions at the same time in this step. This means, you do not have a chance to discuss with other participants.

When you need to make decisions, Window 3 and Window 4 look like below.



(Example 10: A Window 3 and a Window 4, Step 2, Part 3.)

You have 20 seconds to make decisions. The countdown is displayed in the upper right corner of the screen.

When all participants submit their decisions, the computer program with provide feedback on your choice of plans, your neighbours' choices of plans, and your earnings in this round. This round ends here.

In Step 1 of the next round, as we mentioned before, in Window 1, if the computer programme gives you a chance to consider whether to cancel a certain neighbour or not, her choice of plans in Step 2 of this round will appear in Column 2 of Window 1.

(Example 11: A Window 1, Step 1, Part 3.)

If you have any questions, please raise your hand.

This earnings table summarizes your earnings in this round. This table contains 12 sub-tables, from Sub-table A to Sub-table L.

Earnings Table

Sub-table A. My number of neighbours = 0

I do not have to choose plans.
My earnings are 30 RMB.

Sub-table B. My number of neighbours = 1

My Plan / My neighbours' most common Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	48.0	50.0	52.0	54.0
Plan 2	40.0	42.0	44.0	46.0
Plan 3	32.0	34.0	36.0	38.0
Plan 4	24.0	26.0	28.0	30.0

Sub-table C. My number of neighbours = 2

My Plan / My neighbours' most common Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	60.0	63.3	66.7	70.0
Plan 2	46.7	50.0	53.3	56.7
Plan 3	33.3	36.7	40.0	43.3
Plan 4	20.0	23.3	26.7	30.0

Sub-table D. My number of neighbours = 3

My Plan / My neighbours' most common Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	75.0	78.8	82.5	86.3
Plan 2	56.3	60.0	63.8	67.5
Plan 3	37.5	41.3	45.0	48.8
Plan 4	18.8	22.5	26.3	30.0

Sub-table E. My number of neighbours = 4

My Plan / My neighbours' most common Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	93.0	96.8	100.6	104.4
Plan 2	68.2	72.0	75.8	79.6
Plan 3	43.4	47.2	51.0	54.8
Plan 4	18.6	22.4	26.2	30.0

Sub-table F. My number of neighbours = 5

My Plan / My neighbours' most common Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	111.0	114.8	118.7	122.5
Plan 2	80.2	84.0	87.8	91.7
Plan 3	49.3	53.2	57.0	60.8
Plan 4	18.5	22.3	26.2	30.0

Sub-table G. My number of neighbours = 6

My Plan / My neighbours' most common Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	93.0	98.6	104.1	109.7
Plan 2	66.4	72.0	77.6	83.1
Plan 3	39.9	45.4	51.0	56.6
Plan 4	13.3	18.9	24.4	30.0

Sub-table H. My number of neighbours = 7

My Plan / My neighbours' most common Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	75.0	81.9	88.8	95.6
Plan 2	53.1	60.0	66.9	73.8
Plan 3	31.3	38.1	45.0	51.9
Plan 4	9.4	16.3	23.1	30.0

Sub-table I. My number of neighbours = 8

My Plan / My neighbours' most common Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	60.0	67.8	75.6	83.3
Plan 2	42.2	50.0	57.8	65.6
Plan 3	24.4	32.2	40.0	47.8
Plan 4	6.7	14.4	22.2	30.0

Sub-table J. My number of neighbours = 9

My Plan / My neighbours' most common Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	48.0	56.4	64.8	73.2
Plan 2	33.6	42.0	50.4	58.8
Plan 3	19.2	27.6	36.0	44.4
Plan 4	4.8	13.2	21.6	30.0

Sub-table K. My number of neighbours = 10

My Plan / My neighbours' most common Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	42.0	50.7	59.5	68.2
Plan 2	29.3	38.0	46.7	55.5
Plan 3	16.5	25.3	34.0	42.7
Plan 4	3.8	12.5	21.3	30.0

Sub-table L. My number of neighbours = 11

My Plan / My neighbours' most common Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	39.0	47.9	56.8	65.8
Plan 2	27.1	36.0	44.9	53.8
Plan 3	15.2	24.1	33.0	41.9
Plan 4	3.3	12.2	21.1	30.0

Now we introduce how your earnings are determined.

Your earnings depend on 2 conditions.

Condition 1: Your earnings are determined by your plan and the most common plan of all your neighbours.

First, we take Sub-table E (assume you have 4 neighbours) as an example to introduce how to read your earnings in each sub-table according to Condition 1.

① The row title of Sub-table E lists the plans that you can choose.

Sub-table E. My number of neighbours = 4

My Plan / My neighbour's most common Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	93.0	96.8	100.6	104.4
Plan 2	68.2	72.0	75.8	79.6
Plan 3	43.4	47.2	51.0	54.8
Plan 4	18.6	22.4	26.2	30.0

② The column title of Sub-table E lists the possibilities of the most common plan from all your neighbours.

Sub-table E. My number of neighbours = 4

My Plan / My neighbour's most common Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	93.0	96.8	100.6	104.4
Plan 2	68.2	72.0	75.8	79.6
Plan 3	43.4	47.2	51.0	54.8
Plan 4	18.6	22.4	26.2	30.0

By ① and ②: when you have 4 neighbours, your earnings are 47.2 RMB, if

- You chose Plan 2, and
- the most common plan of all your neighbours is Plan 3.

Sub-table E. My number of neighbours = 4

My Plan / My neighbour's most common Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	93.0	96.8	100.6	104.4
Plan 2	68.2	72.0	75.8	79.6
Plan 3	43.4	47.2	51.0	54.8
Plan 4	18.6	22.4	26.2	30.0

Now, we still take Sub-table E as an example, to summarize rules of earnings within each sub-table.

Rule 1: Given your choice, your earnings increase when the most common choice of your neighbours is a **higher Plan**. For example, if your choice is "Plan 2", your earnings are 22.4 RMB when their most common choice is Plan 4, while your earnings increase to 96.8 RMB when their most common choice is Plan 1.

Sub-table E. My number of neighbours = 4

My Plan / My neighbour's most common Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	93.0	96.8	100.6	104.4
Plan 2	68.2	72.0	75.8	79.6
Plan 3	43.4	47.2	51.0	54.8
Plan 4	18.6	22.4	26.2	30.0

Rule 2: Given the most common choice of your neighbours, your earnings increase when you choose a **lower Plan**. For example, if their most common choice is Plan 3, your earnings are 43.4 RMB when your choice is Plan 1, while your earnings increase to 54.8 RMB when your choice is Plan 4.

Sub-table E. My number of neighbours = 4

My Plan / My neighbour's most common Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	93.0	96.8	100.6	104.4
Plan 2	68.2	72.0	75.8	79.6
Plan 3	43.4	47.2	51.0	54.8
Plan 4	18.6	22.4	26.2	30.0

Rule 3: It might be the case that there are 2 or more most common choices that arise by your neighbours. For example, among your 4 neighbours, 2 of them choose Plan 2 and the other 2 choose Plan 4. If it happens, the **lowest plan** will be detected by the computer programme to determine your earnings.
Therefore, your earnings are 22.4 RMB rather than 72.0 RMB when your choice is Plan 2.

Sub-table E. My number of neighbours = 4

My Plan / My neighbour's most common Plan	Plan 1	Plan 2	Plan 3	Plan 4
Plan 1	93.0	96.8	100.6	104.4
~~Plan 2~~	~~68.2~~	~~72.0~~	75.8	79.6
Plan 3	43.4	47.2	51.0	54.8
Plan 4	18.6	22.4	26.2	30.0

Rule 4: Like you, each of your neighbours' earnings depend on her choice and the most common choice of her neighbours**. Since you are also one of the neighbours of each of them**, your choice therefore will also partly determine each of your neighbours' earnings.

Rules 1 to 4 explained above apply to **each** sub-table — from B to L.

Condition 2: Your earnings are determined by the number of neighbours you obtain. Rules of earnings are summarized below.

Rule 5: When you have 0 neighbours, you do not have to choose plans. Your earnings are 30 RMB (Sub-table A).

Rule 6: When you have >=1 neighbours and the most common choice of all your neighbours is Plans 1, 2 or 3:

a. When your number of neighbours increases from 1 to 5, having an extra neighbour leads you to obtain **higher** earnings.
Suppose your choice is Plan 1, and the most common choice of your neighbours is Plan 1 as well. Your earnings increase from 48.0 RMB (having 1 neighbour, Sub-table B) to 111.0 RMB (having 5 neighbours, Sub-table F).

b. When your number of neighbours decreases from 5 to 11, having an extra neighbour leads you to obtain **lower** earnings.
Suppose your choice is Plan 1, and the most common choice of your neighbours is Plan 1 as well. Your earnings decrease from 11.0 RMB (having 6 neighbours, Sub-table F) to 39.0 RMB (having 11 neighbours, Sub-table L).

These choice combinations of earnings are marked as orange .

Rule 7: When you have >=1 neighbours and the most common choice of all your neighbours is Plan 4:

a. When your number of neighbours increases from 1 to 11, having an extra neighbour leads you to obtain **lower** earnings, or remain **unchanged**.
Suppose your choice is Plan 1. Given the most common choice of your neighbours is Plan 4, your earnings will decrease from 24.0 RMB (having 1 neighbour, Sub-table B) to 3.3 RMB (having 11 neighbours, Sub-table L).
Suppose your choice is Plan 4. Given the most common choice of your neighbours is Plan 4, your earnings are always 30 RMB (from Sub-table B to Sub-table L).

These choice combinations of earnings are marked as green .

If you have any questions, please raise your hand.

To confirm your understanding of the instructions, please complete the comprehension questions. The decisions and payoffs used for the questions below are simply for illustrative purposes. Your answers will not affect your earnings in the following tasks.

1. Suppose you have 2 neighbours, and both of your neighbours choose Plan 1. What is your choice of plans that would result in you getting the highest earnings in this case? (D)

   A. Plan 1 B. Plan 2 C. Plan 3 D. Plan 4

2. Suppose you have 7 neighbours, and your choice is Plan 1. In order for you to get the highest earnings in this case, what choice would you need to be the most common choice of your neighbours? (A)

   A. Plan 1 B. Plan 2 C. Plan 3 D. Plan 4

3. Suppose you have 9 neighbours, your choice is Plan 2, 3 of your neighbours choose Plan 1 and 3 of your neighbours choose Plan 2 and 3 of your neighbours choose Plan 3. What are your earnings? (27.6)

4. Suppose you have >=1 neighbours, your choice is Plan 1, and Plan 2 is always the choice of your neighbours. What number of neighbours would result in you reaching the highest earnings in this case? (5)

5. Suppose you have >=1 neighbours, your choice is Plan 3, and Plan 4 is always the choice of your neighbours. What number of neighbours would result in you reaching the highest earnings in this case? (1)

If you have any questions, please raise your hand.

Correct answers and explanations have been shown on the right-hand side of the screen. Please check your responses. If you have any questions, please raise your hand.

Question 1:

D See Sub-table C — When the most common choice of all your neighbours is Plan 1, your earnings by choosing Plan 1 to Plan 4: 60.0, 63.3, 66.7, 70.0. Therefore, you choosing Plan 4 leads you to the highest earnings.

Question 2:

A See Sub-table H — When you choose Plan 1, your earnings resulting from the most common choice of all your neighbours being Plan 1 to Plan 4: 75.0, 53.1, 31.3, 9.4. Therefore, the most common choice being Plan 1 leads you to the highest earnings.

Question 3:

27.6 See Sub-table J — Your earnings depend on your choice of Plan 2, and the lowest most common choice Plan 3. Therefore, your earnings are 27.6 RMB.

Question 4:

5 See all sub-tables. The earnings from the choice combination of "your choice is Plan 1, and the most common choice of all your neighbours is Plan 2" from Sub-tables B to L (number of neighbours from 1 to 11) are 40.0, 46.7, 56.3, 68.2, 80.2 (number of neighbours is 5, Sub-table F), 66.4, 53.1, 42.2, 33.6, 29.3, 27.1. These earnings increase in the number of neighbours at first and then decrease in the number of neighbours; your highest earning appears when you have 5 neighbours (Sub-table F, RMB 80.2), you therefore wish to have 5 neighbours.

Question 5:

1 See all sub-tables. The earnings from the choice combination of "your choice is Plan 3, and the most common choice of all your neighbours is Plan 4" from Sub-tables B to L (number of neighbours from 1 to 11) are 26.7, 26.3, 26.3, 26.2, 24.4, 23.1, 22.2, 21.6, 21.3, 21.1. These earnings decrease in the number of neighbours; your highest earning appears when you have 1 neighbour (Sub-table B, RMB 28.0), you therefore wish to have 1 neighbour.

Your earnings in this experiment are determined as below. At the end of the experiment, we will randomly pick one round in these 25 rounds, and your earnings in this selected round will be paid to you. Since your earnings in every round have an equal probability to be your earnings in this experiment, please make all decisions seriously in each round.

If you have any questions, please raise your hand.

In Step 1 of Round 1, the computer program has already selected 2 neighbours for you, to be your initial state.

They come from the same music group as you. [They come from the different music group from you.] Their unique ID has been displayed on your computer screen.

This means that in Round 1, the computer program has already determined your neighbours in Step 1, you do not need to make decisions at this step; you and the above 2 participants directly start the interaction of Step 2.

From Round 2, and in each subsequent round, both Step 1 and Step 2's decisions are made by yourself.

If you have any questions, please raise your hand.

A2. Treatment effects toward link formation and link deletion

Table A1 and Table A2 show the effect of treatment dummies, High Integration and High Mobility, on the linking rate and the deleting rate (as we mentioned in Section 5.1.2) towards in-group and out-group candidates, respectively. In Table A1, the coefficients of the dummy High Mobility are negative and are statistically significant throughout all models. It implies that subjects are more likely to initiate links in the treatments of Low Mobility, and this effect exists in the whole experiment. In Table A2, the coefficients of the dummy High Mobility are negative and are statistically significant in models (1) to (8) and are statistically significant in models (9) to (10). Therefore, subjects are more likely to delete links in the treatments of Low Mobility, although this effect does not last till the end of the game towards out-group candidates. Table A1 and Table A2 together show that in the treatments of Low Mobility subjects are more active to take the chances to update their relationships. However, as the opportunities of developing new neighbours are low, subjects are restrained to find more same-identity candidates (by deleting out-group neighbours and linking in-group candidates) even if they are active.

Table A1: Treatment effects on link initiating activity.

Dependent variable	**Linking rate towards in-group candidates**					**Linking rate towards out-group candidates**				
	Rounds 2-25	Rounds 6-25	Rounds 10-25	Rounds 14-25	Rounds 18-25	Rounds 2-25	Rounds 6-25	Rounds 10-25	Rounds 14-25	Rounds 18-25
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
High Integration	-0.024	-0.024	-0.023	-0.012	-0.011	-0.024	-0.017	-0.010	-0.004	-0.008
	(0.039)	(0.042)	(0.045)	(0.050)	(0.054)	(0.036)	(0.039)	(0.040)	(0.044)	(0.047)
High Mobility	-0.119***	-0.110***	-0.106**	-0.105**	-0.101*	-0.127***	-0.107**	-0.104**	-0.112**	-0.105**
	(0.036)	(0.040)	(0.042)	(0.047)	(0.050)	(0.036)	(0.039)	(0.040)	(0.043)	(0.047)
Round dummies										
r6-9	-0.199***					-0.138***				
	(0.020)					(0.020)				
r10-13	-0.253***	-0.053***				0.169***	-0.031*			
	(0.018)	(0.014)				(0.021)	(0.016)			
r14-17	-0.293***	-0.093***	-0.040***			-0.219***	-0.080***	-0.050***		
	(0.020)	(0.021)	(0.014)			(0.020)	(0.018)	(0.013)		
r18-21	-0.321***	-0.121***	-0.068***	-0.028		-0.253***	-0.115***	-0.084***	0-.034**	
	(0.024)	(0.021)	(0.019)	(0.017)		(0.023)	(0.017)	(0.018)	(0.013)	
r21-25	-0.320***	-0.121***	-0.067***	0.027	0.001	-0.252***	-0.114***	-0.083***	-0.033**	0.001
	(0.024)	(0.023)	(0.019)	(0.017)	(0.014)	(0.023)	(0.019)	(0.019)	(0.014)	(0.010)
Constant	0.742***	0.538***	0.481***	0.435***	0.405***	0.637***	0.484***	0.448***	0.400***	0.363***
	(0.032)	(0.032)	(0.033)	(0.042)	(0.046)	(0.031)	(0.035)	(0.038)	(0.040)	(0.043)
# of Obs.	7,138	5,940	4,786	3,581	2,377	7,936	6,597	5,288	3,972	2,644
# of Clusters	32	32	32	32	32	32	32	32	32	32
F	37.14	7.38	4.33	1.87	1.37	23.31	10.42	5.29	3.47	1.73
R2 (adj.)	0.074	0.024	0.017	0.014	0.012	0.060	0.026	0.020	0.018	0.015
Prob.>F	0.000	0.000	0.004	0.140	0.269	0.000	0.000	0.001	0.019	0.181

Note: OLS regressions. Dependent variable is the linking rate towards in-group and out-group candidates. High Integration is a dummy variable that equals 1 for treatments O80 and O30 and equals 0 for treatments I80 and I30. High Mobility is a dummy variable that equals 1 for treatments I80 and O80 and equals 0 for treatments I30 and O30. Round 1 is excluded for all models. Round dummies are included; the excluded categories in models (2) and (6) is Round 2-5, in models (3) and (7) are Rounds 2-5 and Rounds 6-9, in models (4) and (8) are Rounds 2-5, Rounds 6-9 and Rounds 10-13, in models (5) and (10) are Rounds 2-5, Rounds 6-9, Rounds 10-13 and Rounds 14-17. Robust standard errors (in parentheses) are clustered at the session level. p < 0.1*, p < 0.05**, p < 0.01***.

Table A2: Treatment effects on link deleting activity.

Dependent variable	**Deleting rate towards in-group candidates**					**Deleting rate towards out-group candidates**				
	Rounds 2-25	Rounds 6-25	Rounds 10-25	Rounds 14-25	Rounds 18-25	Rounds 2-25	Rounds 6-25	Rounds 10-25	Rounds 14-25	Rounds 18-25
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
High Integration	-0.006	-0.005	-0.009	-0.006	-0.014	-0.036	-0.033	-0.039	-0.024	-0.018
	(0.024)	(0.024)	(0.024)	(0.025)	(0.027)	(0.023)	(0.023)	(0.023)	(0.025)	(0.026)
High Mobility	-0.070***	-0.069**	-0.058**	-0.052*	-0.057*	-0.046**	-0.051**	-0.044*	-0.040	-0.037
	(0.025)	(0.026)	(0.027)	(0.028)	(0.032)	(0.021)	(0.022)	(0.023)	(0.025)	(0.027)
Round dummies										
r6-9	-0.066***					-0.063***				
	(0.015)					(0.019)				
r10-13	-0.099***	-0.033**				-0.117***	-0.054**			
	(0.013)	(0.012)				(0.015)	(0.020)			
r14-17	-0.114***	-0.048***	-0.015*			-0.132***	-0.069***	-0.015		
	(0.014)	(0.015)	(0.008)			(0.0156)	(0.019)	(0.014)		
r18-21	-0.119***	0-.053***	-0.020**	-0.005		-0.150***	-0.088***	-0.034**	-0.018**	
	(0.016)	(0.016)	(0.009)	(0.009)		(0.019)	(0.021)	(0.015)	(0.009)	
r21-25	-0.124***	-0.058***	-0.025*	-0.010	-0.005	-0.155***	-0.092***	-0.038***	-0.023*	-0.005
	(0.019)	(0.019)	(0.014)	(0.011)	(0.010)	(0.019)	(0.022)	(0.012)	(0.011)	(0.007)
Constant	0.234	0.169***	0.130***	0.110***	0.113***	0.262***	0.201***	0.146***	0.120***	0.096***
	(0.023)	(0.029)	(0.027)	(0.030)	(0.035)	(0.024)	(0.026)	(0.026)	(0.027)	(0.033)
# of Obs.	6,394	5,436	4,344	3,250	2,179	6,156	5,256	4,241	3,177	2,123
# of Clusters	32	32	32	32	32	32	32	32	32	32
F	13.09	3.33	2.64	1.86	1.49	14.99	6.29	4.87	3.22	0.86
R2 (adj.)	0.042	0.025	0.015	0.012	0.016	0.046	0.030	0.019	0.012	0.009
Prob.>F	0.000	0.012	0.042	0.142	0.236	0.000	0.000	0.002	0.025	0.472

Note: OLS regressions. Dependent variable is the deleting rate towards in-group and out-group candidates. High Integration is a dummy variable that equals 1 for treatments O80 and O30 and equals 0 for treatments I80 and I30. High Mobility is a dummy variable that equals 1 for treatments I80 and O80 and equals 0 for treatments I30 and O30. Round 1 is excluded for all models. Round dummies are included; the excluded categories in models (2) and (6) is Round 2-5, in models (3) and (7) are Rounds 2-5 and Rounds 6-9, in models (4) and (8) are Rounds 2-5, Rounds 6-9 and Rounds 10-13, in models (5) and (10) are Rounds 2-5, Rounds 6-9, Rounds 10-13 and Rounds 14-17. Robust standard errors (in parentheses) are clustered at the session level. $p < 0.1$*, $p < 0.05$**, $p < 0.01$***.

A3. Development of in-group favouritism over the rounds in linking decisions

Table A3: Development of in-group favouritism over the rounds in link formation decisions.

Dependent variable: Link formation (conditional on receiving an opportunity)	Treatment			
	(1)	(2)	(3)	(4)
	I80	I30	O80	O30
Same Identity	1.047***	0.711*	1.141***	1.218***
	(0.125)	(0.407)	(0.179)	(0.337)
Number of Neighbours at r-1	-0.541***	-0.804***	-0.645***	-1.189***
Round Dummies	(0.147)	(0.239)	(0.136)	(0.205)
r6-9	-0.157	-0.388	0.164	-0.204
	(0.196)	(0.324)	(0.195)	(0.378)
r10-13	-0.081	-0.812***	0.074	-0.421
	(0.279)	(0.263)	(0.270)	(0.288)
r14-17	-0.984**	-0.979***	-0.127	-0.499
	(0.404)	(0.238)	(0.239)	(0.390)
r18-21	-1.196***	-1.607***	-0.369***	-0.976**
	(0.323)	(0.450)	(0.262)	(0.437)
r22-25	-1.177***	-1.788***	-0.426***	-1.071***
	(0.264)	(0.371)	(0.283)	(0.344)
Same Identity × Round				
Same Identity × r6-9	-0.492***	-0.393	-0.755***	-0.552
	(0.106)	(0.621)	(0.206)	(0.451)
Same Identity × r10-13	-0.667***	-0.609	-0.990***	-0.990***
	(0.114)	(0.414)	(0.184)	(0.367)
Same Identity × r14-17	-0.433***	-0.668**	-0.905**	-1.318***
	(0.153)	(0.271)	(0.400)	(0.338)
Same Identity × r18-21	-0.388***	-0.681	-0.789**	-0.731*
	(0.128)	(0.486)	(0.350)	(0.384)
Same Identity × r21-25	-0.253*	-0.212	-0.670*	-0.863**
	(0.130)	(0.297)	(0.192)	(0.362)
Constant	1.969***	4.108***	1.925***	5.275***
	(0.359)	(0.760)	(0.289)	(0.654)
Number of obs.	12,670	4,776	12,502	4,630
Number of clusters	8	8	8	8
Log likelihood	-5633.502	-2438.882	-5315.662	-2166.686

Note: Random-effects logistic regressions. In the models, each set of decisions by one participant in one round represents a group of observations. Coefficients are reported, with standard errors in parentheses. Standard errors are clustered at the session level; each treatment has eight clusters. Since there are eight clusters and more than eight predictors, there are not enough degrees of freedom to conduct a Wald test. Data is only included from relationships in which link initiation is allowed. Dependent variable = 1 if a participant initiates a link to that neighbour, conditional on the link being allowed to be changed. Dummy variable Same Identity = 1 if the proposed candidate has the same social identity as the decision maker. Round dummies are included. $p < 0.1$*, $p < 0.05$**, $p < 0.01$***.

Table A4: Development of in-group favouritism over the rounds in link deletion decisions.

Dependent variable: Link deletion (conditional on receiving an opportunity)	Treatment			
	(1)	(2)	(3)	(4)
	I80	I30	O80	O30
Same Identity	-0.947***	0.028	-1.070***	0.050
	(0.196)	(0.346)	(0.224)	(0.410)
Plan r-1	0.104	0.286**	-0.112	0.213*
	(0.066)	(0.137)	(0.166)	(0.117)
Number of Neighbours at r-1	0.071	0.218*	0.061	0.408***
	(0.054)	(0.112)	(0.069)	(0.076)
Round Dummies				
r6-9	-0.968**	0.021	-0.509	-0.361
	(0.480)	(0.452)	(0.427)	(0.361)
r10-13	-0.742**	0.223	-1.857***	-1.468***
	(0.314)	(0.284)	(0.237)	(0.407)
r14-17	-1.192**	-0.152	-1.759***	-1.042***
	(0.574)	(0.361)	(0.459)	(0.312)
r18-21	-1.725***	-0.404	-1.631***	-1.505***
	(0.608)	(0.783)	(0.549)	(0.546)
r21-25	-1.550***	-0.727	-2.090***	-1.499***
	(0.592)	(0.744)	(0.647)	(0.527)
Same Identity × Round				
Same Identity × r6-9	0.396	-0.097	0.594	-0.230
	(0.274)	(0.542)	(0.509)	(0.308)
Same Identity × r10-13	-0.024	-0.560	1.733***	0.398
	(0.429)	(0.366)	(0.547)	(0.518)
Same Identity × r14-17	0.047	-0.736*	1.085**	0.007
	(0.334)	(0.423)	(0.4254)	(0.911)
Same Identity × r18-21	0.289	-0.069	0.841*	0.203
	(0.369)	(0.622)	(0.509)	(0.453)
Same Identity × r21-25	0.112	0.147	1.232***	-0.067
	(0.578)	(0.463)	(0.473)	(0.795)
Constant	-2.113***	-3.553***	-2.432***	-3.971***
	(0.383)	(1.003)	(0.699)	(0.662)
Number of obs.	7,794	2,966	7,960	2,986
Number of clusters	8	8	8	8
Log likelihood	-1799.659	-1125.643	-1557.169	-1001.877

Note: Random-effects logistic regressions. In the models, each set of decisions by one participant in one round represents a group of observations. Coefficients are reported, with standard errors in parentheses. Standard errors are clustered at the session level; each treatment has eight clusters. Since there are eight clusters and more than eight predictors, there are not enough degrees of freedom to conduct a Wald test. Data is only included from relationships in which link deletion is allowed. Dependent variable = 1 if a participant deletes a link with that neighbour. Dummy variable Same Identity = 1 if the neighbour has the same social identity as the decision maker. Plan r-1 represents the choice of Plan 1, 2, 3 or 4 (from the best to the worst) of the neighbour at r-1. Round dummies are included. $p < 0.1$*, $p < 0.05$**, $p < 0.01$***.

A4. Decision differences across treatments

Table A5: Decisions differences across treatments.

Dependent variable:	**Link formation (conditional on receiving an opportunity)**		**Link deletion (conditional on receiving an opportunity)**		
	(1)	(2)	(3)	(4)	(5)
Same Identity	0.506***	0.369***	-0.369***	-0.355**	-0.368***
	(0.067)	(0.130)	(0.113)	(0.160)	(0.113)
Plan r-1			0.115	0.114	0.286**
			(0.072)	(0.071)	(0.123)
Number of Neighbours at r-1	-0.664***	-0.663***	0.133***	0.130***	0.136***
	(0.096)	(0.096)	(0.042)	(0.042)	(0.041)
High Integration	-0.029	-0.039	-0.344	-0.550*	--0.096
	(0.198)	(0.215)	(0.315)	(0.299)	(0.505)
High Mobility	-0.721***	-0.821***	-0.620**	-0.426	-0.187
	(0.204)	(0.220)	(0.313)	(0.304)	(0.483)
Same Identity × High Integration		0.022		0.425**	
		(0.127)		(0.191)	
Same Identity × High Mobility		0.231*		-0.403**	
		(0.139)		(0.177)	
Plan r-1 × High Integration					-0.114
					(0.134)
Plan r-1 × High Mobility					-0.202
					(0.130)
Round dummies					
r6-9	-0.303***	-0.399**		-0.407**	-0.399**
	(0.120)	(0.190)		(0.190)	(0.191)
r10-13	-0.484***	-0.764***		-0.775***	-0.764***
	(0.133)	(0.150)		(0.152)	(0.150)
r14-17	-0.898***	-1.021***		-1.031***	-1.026***
	(0.188)	(0.239)		(0.241)	(0.241)
r18-21	-1.154***	-1.197***		-1.206***	-1.211***
	(0.192)	(0.248)		(0.249)	(0.247)
r21-25	-1.152***	-1.302***		-1.309***	-1.318***
	(0.159)	(0.292)		(0.293)	(0.291)
Constant	3.257***	3.315***	-2.263***	-2.235***	-2.665***
	(0.074)	(0.350)	(0.365)	(0.361)	(0.473)
Number of obs.	34,578	34,578	21,706	21,706	21,706
Number of clusters	32	32	32	32	32
Wald chi2	354.34	404.07	63.90	109.32	64.29
Log likelihood	-15734.233	-15733.041	-5549.673	-5546.585	-5546.585

Note: Random-effects logistic regressions. In the models, each set of decisions by one participant in one round represents a group of observations. Coefficients are reported, with standard errors in parentheses. Standard errors are clustered at the session level, the pooled data has 32 clusters. Data is only included from relationships in which either link initiation (column (1) and (2)) or link deletion (columns (3) to (5)) is allowed. Dependent variable = 1 if a participant initiates (column (1) and (2)) or deletes a link (columns (3) to (5)) with that candidate. Dummy variable Same Identity = 1 if the neighbour has the same social identity as the decision maker. Plan r-1 represents the choice of Plan 1, 2, 3 or 4 (from the best to the worst) of the neighbour at r-1. High Integration is a dummy variable that equals 1 for treatments O80 and O30 and equals 0 for treatments I80 and I30. High Mobility is a dummy variable that equals 1 for treatments I80 and O80 and equals 0 for treatments I30 and O30. Round dummies are included. p < 0.1*, p < 0.05**, p < 0.01***.

A5. Do subjects reward and punish in-group and out-group members differently?

Table A6: Effects of contributions and identity on link deletion decisions.

Dependent variable: Link deletion (conditional on receiving an opportunity)	Treatment			
	(1)	(2)	(3)	(4)
	I80	I30	O80	O30
Same Identity	-1.264***	-0.214	-0.602	0.450
	(0.393)	(0.386)	(0.428)	(0.459)
Plan r-1	-0.010	0.282**	-0.154	0.291**
	(0.121)	(0.134)	(0.190)	(0.115)
Number of Neighbours at r-1	0.076	0.220**	0.035	0.403***
	(0.052)	(0.106)	(0.073)	(0.073)
Round dummies				
r6-9	-0.751*	-0.077	-0.244	-0.458
	(0.421)	(0.150)	(0.389)	(0.347)
r10-13	-0.750***	-0.082	-1.035***	-1.261***
	(0.287)	(0.192)	(0.215)	(0.305)
r14-17	-1.163*	-0.550	-1.264***	-1.042***
	(0.617)	(0.348)	(0.318)	(0.210)
r18-21	-1.565***	-0.473	-1.255***	-1.409***
	(0.569)	(0.527)	(0.336)	(0.368)
r21-25	-1.488**	-0.691	-1.526***	-1.524***
	(0.628)	(0.632)	(0.447)	(0.420)
Same Identity × Plan r-1	0.210	0.008	0.122	-0.181
	(0.161)	(0.113)	(0.175)	(0.219)
Constant	-1.955***	-3.390***	-2.515***	-4.120***
	(0.511)	(0.796)	(0.682)	(0.608)
Number of obs.	7,794	2,966	7,960	2,986
Number of clusters	8	8	8	8
Log likelihood	-1799.277	-1128.1757	-1566.434	-1002.004

Note: Random-effects logistic regressions. In the models, each set of decisions by one participant in one round represents a group of observations. Coefficients are reported, with standard errors in parentheses. Standard errors are clustered at the session level; each treatment has eight clusters. Since there are eight clusters and more than eight predictors, there are not enough degrees of freedom to conduct a Wald test. Data is only included from relationships in which link deletion is allowed. Dependent variable = 1 if a participant deletes a link with that neighbour. Dummy variable Same Identity = 1 if the neighbour has the same social identity as the decision maker. Plan r-1 represents the choice of Plan 1, 2, 3 or 4 (from the best to the worst) of the neighbour at r-1. Round dummies are included. $p < 0.1$*, $p < 0.05$**, $p < 0.01$***.

A6. Are same-identity candidates more likely to be allowed by the computer programme to be added (deleted) in link formation (link deletion) decisions?

Table A7: Effects of luck on link formation decisions.

Dependent variable: Allowed to update relationship	Treatment (or pooled data)				
	(1)	(2)	(3)	(4)	(5)
	I80	I30	O80	O30	Pooled data
Same Identity	-0.050	-0.035	0.010	-0.023	-0.025
	(0.041)	(0.053)	(0.038)	(0.024)	(0.020)
High Integration					-0.015
					(0.027)
High Mobility					2.341***
					(0.025)
Round dummies					
r6-9	0.038	0.229**	-0.011	0.155*	0.115**
	(0.095)	(0.089)	(0.108)	(0.093)	(0.047)
r10-13	0.121	0.406***	0.149	0.293***	0.257***
	(0.082)	(0.041)	(0.119)	(0.094)	(0.044)
r14-17	0.098	0.425***	0.183	0.350***	0.281***
	(0.101)	(0.066)	(0.122)	(0.110)	(0.051)
r18-21	0.122	0.433***	0.043	0.275***	0.238***
	(0.092)	(0.096)	(0.142)	(0.073)	(0.054)
r21-25	0.197	0.414***	0.087	0.315***	0.270***
	(0.091)	(0.107)	(0.065)	(0.088)	(0.048)
Constant	1.345**	-1.183***	1.368***	-1.146***	-1.080***
	(0.079)	(0.063)	(0.076)	(0.054)	(0.040)
Number of obs.	15,734	16,474	15,470	16,422	64,100
Number of clusters	8	8	8	8	32
Wald chi2	130.20	260.24	18.35	61.97	10218.83
Prob > chi2	0.000	0.000	0.005	0.000	0.000
Log likelihood	-7752.802	-9882.201	-7556.246	-9742.396	-34953.631

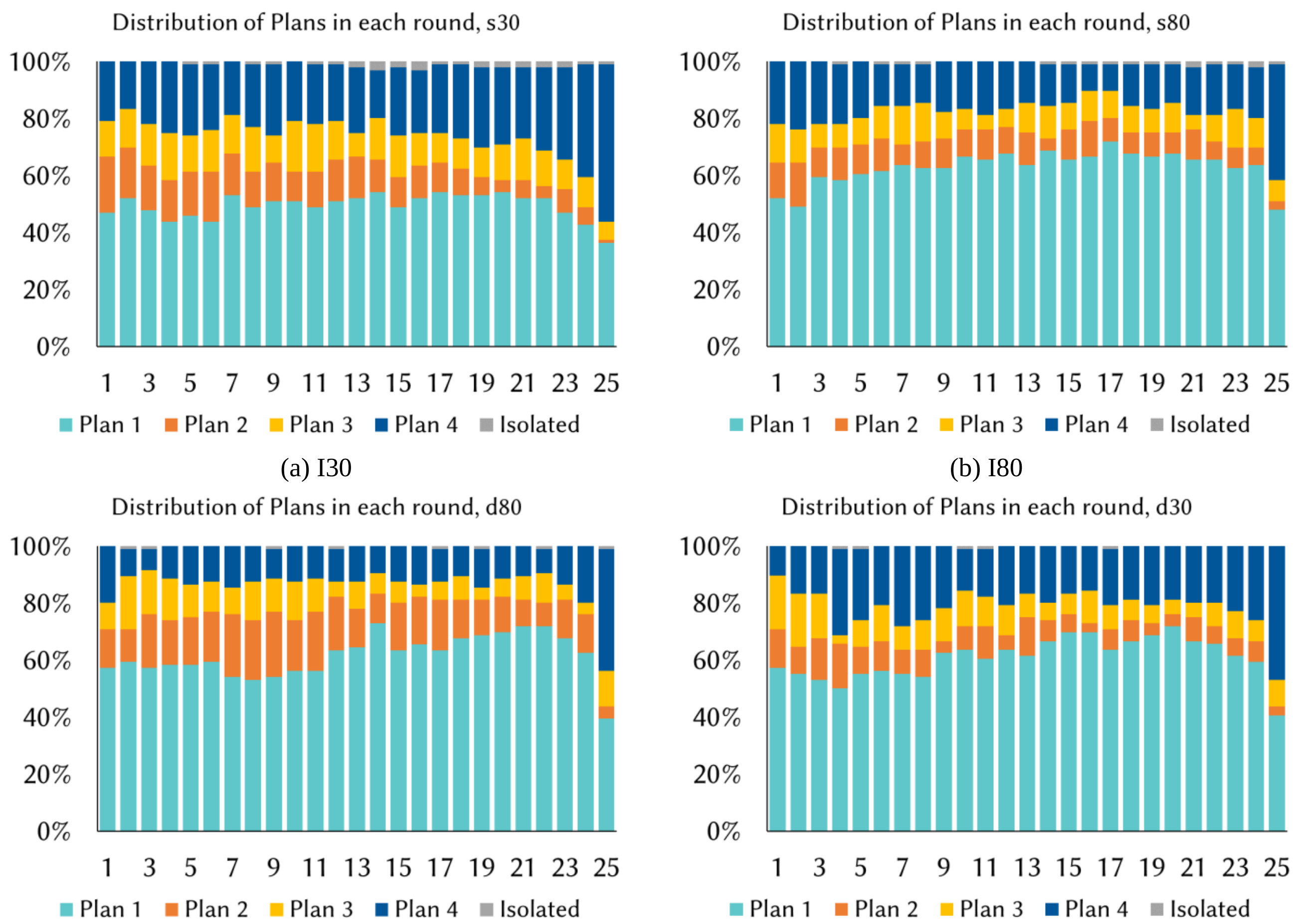
Note: Random-effects logistic regressions. In the models, each set of decisions by one participant in one round represents a group of observations. Coefficients are reported, with standard errors in parentheses. Standard errors are clustered at the session level; each treatment has eight clusters (in models (1) to (4)) and the pooled data has 32 clusters (in model (5)). Dependent variable = 1 if the computer allows this link to be added. Dummy variable Same Identity = 1 if the neighbour has the same social identity as the decision maker. High Integration is a dummy variable that equals 1 for treatments O80 and O30 and equals 0 for treatments I80 and I30. High Mobility is a dummy variable that equals 1 for treatments I80 and O80 and equals 0 for treatments I30 and O30. Round dummies are included. $p < 0.1$*, $p < 0.05$**, $p < 0.01$***.

Table A8: Effects of luck on link deletion decisions.

Dependent variable: Allowed to update relationship	Treatment (or pooled data)				
	(1)	(2)	(3)	(4)	(5)
	I80	I30	O80	O30	Pooled data
Same Identity	-0.098*	-0.071	0.090	0.114*	0.015
	(0.060)	(0.064)	(0.062)	(0.066)	(0.034)
High Integration					-0.016
					(0.031)
High Mobility					2.127***
					(0.034)
Round dummies					
r6-9	-0.455***	-1.246***	-0.615***	-0.781***	-0.799***
	(0.171)	(0.166)	(0.079)	(0.087)	(0.089)
r10-13	-0.624***	-1.388***	-0.325**	-1.017***	-0.876***
	(0.223)	(0.107)	(0.157)	(0.112)	(0.104)
r14-17	-0.506**	-1.430***	-0.467***	-1.189***	-0.939***
	(0.217)	(0.080)	(0.093)	(0.142)	(0.103)
r18-21	-0.458***	-1.377***	-0.431**	-1.066***	-0.873***
	(0.141)	(0.132)	(0.211)	(0.168)	(0.111)
r21-25	-0.448**	-1.366***	-0.398***	-1.212***	-0.903***
	(0.184)	(0.120)	(0.073)	(0.110)	(0.102)
Constant	1.980***	0.599***	1.814***	0.222**	0.117***
	(0.180)	(0.092)	(0.133）	(0.103)	(0.090)
Number of obs.	9,584	8,830	9,851	8,881	37,146
Number of clusters	8	8	8	8	32
Wald chi2	45.76	684.11	254.93	228.21	4053.72
Prob > chi2	0.000	0.000	0.000	0.000	0.000
Log likelihood	-4584.810	-5487.667	-4779.767	-5553.307	-20482.934

Note: Random-effects logistic regressions. In the models, each set of decisions by one participant in one round represents a group of observations. Coefficients are reported, with standard errors in parentheses. Standard errors are clustered at the session level; each treatment has eight clusters (in models (1) to (4)) and the pooled data has 32 clusters (in model (5)). Dependent variable = 1 if the computer allows this link to be deleted. Dummy variable Same Identity = 1 if the neighbour has the same social identity as the decision maker. High Integration is a dummy variable that equals 1 for treatments O80 and O30 and equals 0 for treatments I80 and I30. High Mobility is a dummy variable that equals 1 for treatments I80 and O80 and equals 0 for treatments I30 and O30. Round dummies are included. p < 0.1*, p < 0.05**, p < 0.01***.

A7. Distribution of Plans by round and by treatment



(a) I30 (b) I80

(c) O30 (d) O80

Figure A8: Distribution of plans over the rounds, by treatment.

A8. Effects of proportion of same-identity neighbours at the current round on contribution decision

Table A10: Effects of proportion of same-identity neighbours on choice of plans.

	Rounds				
Dependent variable:	(1)	(2)	(3)	(4)	(5)
Choice of plans of subjects	Rounds 2-25	Rounds 6-25	Rounds 10-25	Rounds 14-25	Rounds 18-25
High Integration	-0.176	-0.175	-0.215	-0.234	-0.265
	(0.156)	(0.153)	(0.154)	(0.159)	(0.165)
High Mobility	-0.230	-0.242	-0.240	-0.263	-0.272
	(0.154)	(0.151)	(0.153)	(0.156)	(0.164)
Proportion of same-identity neighbours at r	-0.344**	-0.352*	-0.452**	-0.440**	-0.506**
	(0.166)	(0.190)	(0.202)	(0.214)	(0.225)
Round dummies					
r6-9	-0.041				
	(0.037)				
r10-13	-0.139**	-0.098**			
	(0.053)	(0.042)			
r14-17	-0.205***	-0.164**	-0.065*		
	(0.069)	(0.062)	(0.037)		
r18-21	0.177**	-0.136*	-0.037	0.028	
	(0.077)	(0.073)	(0.048)	(0.030)	
r21-25	0.121	0.162**	0.260***	0.326***	0.297***
	(0.079)	(0.073)	(0.049)	(0.038)	(0.033)
Constant	2.356***	2.323***	2.296***	2.245	2.328***
	(0.195)	(0.195)	(0.189)	(0.194)	(0.212)
# of Obs.	9,155	7,625	6,096	4,567	3,046
# of Clusters	32	32	32	32	32
F	17.32	19.16	22.03	25.99	29.71
R2 (adj.)	0.027	0.029	0.036	0.041	0.044
Prob.>F	0.000	0.000	0.000	0.000	0.000

Note: OLS regressions. Dependent variable is plan chosen by subjects. Independent variable is the proportion of neighbours with the same identity as each subject. High Integration is a dummy variable such that it equals 1 for treatment O80 and O30 and it equals 0 for treatment I80 and I30. High Mobility is a dummy variable such that it equals 1 for treatments I80 and O80 and it equals 0 for treatments I30 and O30. Round 1 is excluded for all models. Round dummies are included; the omitted category in model (1) is Round 2-5, in model (2) is Round 6-9, in model (3) is Round 10-13, in model (4) is Round 14-17 and in model (5) is Round 18-21. Robust standard errors are clustered at the session level. $p < 0.1$*, $p < 0.05$**, $p < 0.01$***.

Table A11: Effects of proportion of same-identity neighbours on choice of plan, by treatment.

Dependent variable: Choice of plans of subjects	Treatment			
	(1)	(2)	(3)	(4)
	I80	I30	O80	O30
Proportion of same-identity neighbours at r	-0.586**	-1.001**	0.180	0.241
	(0.245)	(0.285)	(0.304)	(0.270)
Round dummies				
r6-9	-0.169	-0.110**	0.036	-0.026
	(0.090)	(0.041)	(0.089)	(0.067)
r10-13	-0.236**	-0.157*	-0.027	-0.218
	(0.093)	(0.074)	(0.108)	(0.151)
r14-17	-0.324**	-0.169**	-0.132	-0.277
	(0.096)	(0.068)	(0.127)	(0.216)
r18-21	-0.257**	-0.093	-0.166	-0.277
	(0.106)	(0.094)	(0.126)	(0.238)
r21-25	-0.003	0.222***	0.134	0.056
	(0.133)	(0.048)	(0.133)	(0.241)
Constant	2.327***	2.713***	1.678***	1.934***
	(0.200)	(0.261)	(0.278)	(0.268)
# of Obs.	2,286	2,274	2,296	2,299
# of Clusters	8	8	8	8
F	63.60	16.34	94.16	18.73
R2 (adj.)	0.024	0.042	0.010	0.013
Prob.>F	0.000	0.001	0.000	0.001

Note: OLS regressions. Dependent variable is plan chosen by subjects. Independent variable is the proportion of neighbours with the same identity as each subject. Round dummies are included. Robust standard errors are clustered at the session level; each treatment has eight clusters. Round 1 is excluded for all models. $p < 0.1$*, $p < 0.05$**, $p < 0.01$***.

A9. Treatment effects on payoffs

Table A12: Treatment effects on payoffs.

Dependent variable: Payoff	Rounds				
	(1)	(2)	(3)	(4)	(5)
	Rounds 2-25	Rounds 6-25	Rounds 10-25	Rounds 14-25	Rounds 18-25
High Integration	6.432	6.748	7.853	8.444	8.402
	(6.502)	(6.614)	(6.660)	(6.875)	(6.946)
High Mobility	6.902	7.246	7.610	8.419	9.129
	(6.502)	(6.614)	(6.660)	(6.875)	(6.946)
Round dummies					
r6-9	6.278***				
	(1.412)				
r10-13	9.812***	3.534**			
	(1.631)	(1.645)			
r14-17	11.383***	5.104**	1.571		
	(2.186)	(2.186)	(1.230)		
r18-21	12.390***	6.112**	2.578	1.008	
	(2.505)	(2.497)	(1.600)	(0.780)	
r21-25	3.876***	-2.403	-5.936***	-7.507	-8.514
	(2.697)	(2.752)	(1.803)	(1.360)	(1.188)
Constant	67.411***	73.360***	76.159***	77.029***	77.703
	(6.707)	(6.907)	(6.795)	(7.034)	(7.177)
# of Obs.	9,216	7,680	6,144	4,608	3,072
# of Clusters	32	32	32	0.000	0.000
F	16.92	15.83	15.94	19.95	27.04
R2 (adj.)	0.037	0.036	0.036	0.043	0.044
Prob.>F	0.000	0.000	0.000	0.000	0.000

Note: OLS regressions. Dependent variable is payoff of subjects. High Integration is a dummy variable such that it equals 1 for treatment O80 and O30 and it equals 0 for treatment I80 and I30. High Mobility is a dummy variable such that it equals 1 for treatments I80 and O80 and it equals 0 for treatments I30 and O30. Round 1 is excluded for all models. Round dummies are included; the omitted category in model (1) is Round 2-5, in model (2) is Round 6-9, in model (3) is Round 10-13, in model (4) is Round 14-17 and in model (5) is Round 18-21. Robust standard errors are clustered at the session level. $p < 0.1$*, $p < 0.05$**, $p < 0.01$***.

A10. Treatment effects on the modal plan

Table A13: Treatment effects on modal plan.

Dependent variable: **Most common planc of neighbours**	Rounds				
	(1)	(2)	(3)	(4)	(5)
	Rounds 2-25	Rounds 6-25	Rounds 10-25	Rounds 14-25	Rounds 18-25
High Integration	0.239	-0.257	0.307	-0.337	-0.335
	(0.234)	(0.231)	(0.231)	(0.237)	(0.245)
High Mobility	0.276	-0.276	-0.283	-0.303	-0.308
	(0.234)	(0.231)	(0.231)	(0.237)	(0.245)
Round dummies					
r6-9	-0.105*				
	(0.053)				
r10-13	-0.210***	0.105			
	(0.074)	(0.068)			
r14-17	-0.271**	-0.165	-0.061		
	(0.099)	(0.095)	(0.054)		
r18-21	-0.292**	-0.187*	-0.082	-0.021	
	(0.108)	(0.101)	(0.063)	(0.041)	
r21-25	0.091	0.197*	0.301***	0.362***	0.383***
	(0.111)	(0.106)	(0.066)	(0.052)	(0.046)
Constant	2.186***	2.089***	2.013***	1.978***	0.958***
	(0.260)	(0.260)	(0.246)	(0.251)	(0.264)
# of Obs.	9,216	7,680	6,144	4,608	3,072
# of Clusters	32	32	32	32	32
F	17.17	19.45	22.64	28.60	37.47
R2 (adj.)	0.037	0.039	0.047	0.056	0.056
Prob.>F	0.000	0.000	0.000	0.000	0.000

Note: OLS regressions. Dependent variable is most common choice of plan of a subject's neighbours. High Integration is a dummy variable such that it equals 1 for treatment O80 and O30 and it equals 0 for treatment I80 and I30. High Mobility is a dummy variable such that it equals 1 for treatments I80 and O80 and it equals 0 for treatments I30 and O30. Round 1 is excluded for all models. Round dummies are included the omitted category in model (1) is Round 2-5, in model (2) is Round 6-9, in model (3) is Round 10-13, in model (4) is Round 14-17 and in model (5) is Round 18-21. Robust standard errors are clustered at the session level. $p < 0.1$*, $p < 0.05$**, $p < 0.01$***.

A11. Effects of historical contributions on link deletion decisions over rounds

Table A14: Effects of historical contributions on link deletion decisions.

Dependent variable: Link deletion (conditional on receiving an opportunity)	Treatment			
	(1)	(2)	(3)	(4)
	I80	I30	O80	O30
Same Identity	-0.822***	-0.179	-0.376*	0.060
	(0.187)	(0.183)	(0.226)	(0.185)
Plan r-1	-0.291**	0.266*	-0.301***	-0.108
	(0.128)	(0.101)	(0.114)	(0.238)
Number of Neighbours at r-1	0.090*	-0.270**	0.039	0.406***
	(0.053)	(0.145)	(0.069)	(0.063)
Round dummies				
r6-9	-1.787***	-0.714*	-0.811	-1.568**
	(0.559)	(0.372)	(0.725)	(0.668)
r10-13	-1.595***	-1.358***	-1.375*	-1.760***
	(0.510)	(0.319)	(0.718)	(0.601)
r14-17	-2.451***	-2.334***	-1.730***	-0.906*
	(0.666)	(0.427)	(0.640)	(0.465)
r18-21	-3.028***	-3.219***	-1.984***	-2.951***
	(1.113)	(0.644)	(0.548)	(0.571)
r21-25	-3.009***	-3.032***	-1.974**	-2.330**
	(0.987)	(0.532)	(0.971)	(0.958)
Plan r-1 × Round dummies				
Plan r-1 × r6-9	0.496***	0.276	0.291	0.531**
	(0.179)	(0.219)	(0.220)	(0.308)
Plan r-1 × r10-13	0.406**	0.560***	0.181	0.240
	(0.183)	(0.130)	(0.270)	(0.344)
Plan r-1 × r14-17	0.627***	0.767***	0.248	-0.151
	(0.133)	(0.092)	(0.204)	(0.284)
Plan r-1 × r18-21	0.691**	1.099***	0.393**	0.769***
	(0.299)	(0.157)	(0.167)	(0.224)
Plan r-1 × r21-25	0.708***	0.936***	0.241	0.396
	(0.206)	(0.228)	(0.290)	(0.313
Constant	-1.412***	-2.333***	-2.234***	-3.319***
	(0.528)	(0.780)	(0.585)	(0.645)
Number of obs.	7,794	2,966	7,960	2,986
Number of clusters	8	8	8	8
Log likelihood	-1784.806	-1104.760	-1564.408	-991.667

Note: Random-effects logistic regressions. In the models, each set of decisions by one participant in one round represents a group of observations. Coefficients are reported, with standard errors in parentheses. Standard errors are clustered at the session level; each treatment has eight clusters. Since there are eight clusters and more than eight predictors, there are not enough degrees of freedom to conduct a Wald test. Data is only included from relationships in which link deletion is allowed. Dependent variable = 1 if a participant deletes a link with that neighbour. Dummy variable Same Identity = 1 if the neighbour has the same social identity as the decision maker. Plan r-1 represents the choice of Plan 1, 2, 3 or 4 (from the best to the worst) of the neighbour at r-1. Round dummies are included. $p < 0.1$*, $p < 0.05$**, $p < 0.01$***.