

Rezaei, Sarah; Rosenkranz, Stephanie; Weitzel, Utz; Westbrock, Bastian

Conference Paper

Social Preferences on Networks

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2022: Big Data in Economics

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Rezaei, Sarah; Rosenkranz, Stephanie; Weitzel, Utz; Westbrock, Bastian (2022) : Social Preferences on Networks, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2022: Big Data in Economics, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/264063>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Social Preferences on Networks

Sarah Rezaei* Stephanie Rosenkranz[†] Utz Weitzel[‡]
Bastian Westbrock[§]

July 24, 2021

This study belongs to the research program "Cooperation in Social and Economic Relations", which enjoys a waiver from Utrecht University's Institutional Review Board (IRB). Approval number: FETC17-028.

*Innsbruck University, Department of Economics, Sarah.Rezaei@uibk.ac.at

[†]Utrecht University School of Economics, S.Rosenkranz@uu.nl

[‡]Vrije Universiteit Amsterdam, & Tinbergen Institute & Radboud University, Institute for Management Research, u.weitzel@vu.nl

[§]Fernuniversität Hagen, Institute for Economics, bastian.westbrock@gmail.com

Abstract

Social preferences are a powerful determinant of human behavior. We theoretically and experimentally study their impact on behavior in a local public goods game on a fixed network structure. The key feature of the game is that it has multiple equilibria which differ widely in terms of their payoff consequences. This makes coordination a key challenge. The coordination problem is exacerbated by the fact that socially concerned players might disagree about the “right” payoff ordering. However, we show that when players’ social preferences “fit” the network positions they occupy, players successfully coordinate on a very fine-grained equilibrium set. How easily the preference requirements are met depends on a property of the network structure: neighborhood nestedness. This means that equilibrium selection succeeds in small, tightly connected structures and in very centralized networks, but it fails in loosely-connected local interaction structures. We test our predictions in an experiment, which gives players ample opportunities for coordination and randomly allocates players to networks that widely differ in terms of their degree of nestedness. All our predictions are confirmed.

JEL: D85, C70, C91, H41

Keywords: social preferences, network games, equilibrium selection, network nestedness

1 Introduction

In our daily lives, we are involved in many social interactions and constantly struggle to divide our time, effort, and resources with others. These interactions can oftentimes be viewed as public goods games where each partner's investment is enjoyed by everybody else involved in the interaction. Examples are gardening efforts which also our neighbors enjoy, the preventive measures we take in a pandemic to protect our contacts or the time we dedicate to a joint project with our co-workers. Many experimental and empirical studies suggest that social preferences shape our behavior in such public goods situations.¹ Yet, it is not clear how social preferences play out in a network of interdependent public goods situations.

Studying this topic in the field is difficult because of many obstacles, starting with the difficulty of collecting the required data and ending with the privacy concerns it would raise. In this paper, we therefore study it theoretically and experimentally. In particular, we extend on the seminal game by Bramoullé and Kranton (2007) which has much in common with the social dilemma described above. Players are embedded in a fixed network, and they make investments in a local public good which they share with their direct network neighbors. Unlike in a standard public goods game, the benefits of this investment are not linear. Instead, there is a strictly positive investment level that even a payoff maximizer would want to invest in, and players aim to fill the gap between this optimum and their neighbors' investments. This strategic substitutes property speaks to many of the public goods situations described above, as it is oftentimes clear that a public good must be provided, and the only important question is who is going to do it. But strategic substitutability also implies that the Bramoullé and Kranton (2007) game has multiple equilibria because every investment profile that satisfies the threshold condition is an equilibrium.

Three important questions are emerging from here: Do social preferences help players to maintain investments in the game that exceeds the privately optimal level and that are closer to what is required for an efficient level of the public good? Do social preferences help to coordinate behaviour on a unique equilibrium? And if so, do they help to coordinate on equitable payoff distributions when the network structure itself is asymmetric?

To structure our thoughts on these questions, we first develop a theory of social preferences in network games. Towards this end, we incorporate an n -player extension of the Charness and Rabin (2002) social preference function into the Bramoullé and Kranton (2007) game. A nice feature of this function is that it captures several distinct social preference types which real people have been shown to care about, such as altruism, social welfare,

¹For example, a large number of experimental studies has shown that individuals invest more than what is optimal from a pure individual perspective in a public goods game. See, for instance, Eckel and Harwell (2015) for a replication study of four classic experiments.

inequity aversion, competitiveness, and spite.² We then characterize the equilibria of the modified game, which we refer to as the other-regarding equilibria (ORE), where each player has one of these preferences.

Our main result is that, regardless of the exact preference type, a socially concerned player strives for a certain payoff ordering in her network neighborhood. In other words, she chooses an investment level considering the relative payoff consequences for herself and for her neighbors in a way that is consistent with her preference. This strive for a payoff ordering has some important implications for the structure of equilibria. In particular, it implies that many of the equilibria that were maintained in the original game with pure payoff maximizers can be ruled out. In an ORE on the star network, for example, the central player must earn more than at least one of the peripheral players. And, if all players share the same connections, such as in the complete network, they must earn the same. Even more so, when the weights on players' social preference components are small, we show that the set of ORE must even be a strict subset of the payoff-maximizing equilibria on many network structures. In other words, our theory suggests that social preferences facilitate equilibrium selection in the Bramoullé and Kranton (2007) game.

Our theory also suggests, however, that equilibrium selection does not work on all networks and for all preference types alike. Rather, it is tied to two conditions. First, players need, what we call, a set of compatible social preferences, that is, preferences that “fit” the position they occupy in a network. Preference compatibility is, for instance, satisfied in a homogeneous group of players who share the same social preference. It is violated, though, when an altruist is in a central network position or when competitive players occupy the peripheral positions. For such players, our theory predicts an ORE set that is wider than the payoff-maximizing equilibrium set.

The second condition for equilibrium selection is that the neighborhoods of a network need to be nested.³ Nestedness is a well-documented topology of many ecological systems (Mariani, Ren, Bascompte, and Tessone, 2019) and it emerges as the outcome of many evolving social networks (König, Tessone, and Zenou, 2014; Belhaj, Bervoets, and Deroïan, 2016). In our theory, it ensures that socially concerned players can enforce the payoff ordering they desire to maintain in their neighborhood. This means that equilibrium selection works particularly well in tightly connected network structures, such as the complete network, or in highly centralized star-like structures. Multiple equilibria may arise, in contrast, when the network has a loosely-connected local interaction structure, such as the circle network.

²See Bellemare, Kröger, and Van Soest (2008), Falk, Becker, Dohmen, Enke, Huffman, and Sunde (2018), and Kerschbamer and Müller (2020) for empirical evidence on the diversity of social preferences.

³The neighborhoods of two players are nested if the neighborhood of one player is contained in the neighborhood of the other (Mariani, Ren, Bascompte, and Tessone, 2019).

In the second part of our paper, we validate the central mechanisms and predictions of our theory. Our experiment has a number of design features to facilitate the test. First, subjects play the Bramoullé and Kranton (2007) game on a set of networks that differ widely in terms of their degree of nestedness and the organization of nested neighborhoods. Second, our experimental games feature a large strategy space to allow for the full set of payoff-maximizing equilibria as well as for deviations thereof to emerge. Third, to ensure that subjects play equilibria at all, we let subjects adjust their choices in continuous-time with random termination.

To give an outlook on our findings, first, we do not find any indication that social preferences would help to reach efficient levels of the public good. Rather, the majority of investments in our experiment were indistinguishable from payoff-maximizing behavior. If at all, subjects underinvested compared to what a payoff maximizer would invest their stead. Second, we find that subjects' social preferences helped them to coordinate their choices so as to reach the narrow set of ORE predicted by our theory. For the identification of this effect, we make use of our theoretical prediction that a group of social players only converges to a narrow set of equilibria if the social preferences of all group members are compatible. We thus compare the successful coordination attempts of subject groups with and without the proper preference combination. Finally, our experiment suggests that subjects do not make any attempt to reach equitable payoff distributions at all costs. Rather, as predicted by our theory, the majority of investments just reinforced the inequality that was already inherent in a network structure.

In the next section, we relate our contribution to the literature. Section 3 briefly summarizes the Bramoullé and Kranton (2007) game and presents our theoretical predictions for socially concerned players. Section 4 describes the experiment, and Section 5 the findings. The proofs of all our statements, additional evidence from the experiment, and the replication instructions can be found in the appendix.

2 Related literature

Our study relates to the literature on social networks and social preferences. In the networks literature, there are a few other theories with socially concerned players, notably Ghiglino and Goyal (2010), Immorlica, Kranton, Manea, and Stoddard (2017), Bourlès, Bramoullé, and Perez-Richet (2017), and Richefort (2018). A major difference between these theories and ours is that they look at contexts without any strategic interaction between players if it were not for their social comparison concerns. The examples they think of are anonymous market interactions, financial transfers between family members, or one's status in a larger neighborhood. The situation we look at, in contrast, is a complex local interaction game with multiple equilibria.

So, while social preferences merely “shift” the unique equilibrium point in these earlier theories, the players in our game are confronted with the additional problem of how to coordinate their choices.⁴

Our result that social preferences facilitate coordination in such a game is related to one other important branch in the networks literature that aims to tackle the pervasive problem of equilibrium multiplicity. Bramoullé, Kranton, and D’Amours (2014) and Allouch (2015) make clear when the problem is most severe, namely in games where players’ investments are strategic substitutes. Several equilibrium refinement concepts have been proposed so far. Bramoullé and Kranton (2007), for example, study Nash tâtonnement stability, Boncinelli and Pin (2012) stochastic stability, and Galeotti, Goyal, Jackson, Vega-Redondo, and Yariv (2010) look at the coordinating role of the limitation of agents’ information about the network structure beyond their immediate neighborhood. All these concepts select roughly the same type of equilibrium on the class of asymmetric networks: the periphery-sponsored equilibrium where players with fewer connections invest more than players with more connections. Social preferences in our theory help to select just the same type of equilibrium. Moreover, however, they also explain why individuals typically succeed to coordinate on equal-split equilibria on a 2-player dyad network and fail to coordinate on a circle network—phenomenon the previous concepts have nothing to say about.

Regarding the experimental networks literature, most existing studies test equilibrium selection in the lab, with mixed evidence. Charness, Feri, Meléndez-Jiménez, and Sutter (2014), for instance, investigate the role of uncertainty about the network structure. Yet, they conclude that uncertainty does not facilitate coordination per se because equilibrium play is also very high in treatments with complete information. The guiding principle to equilibrium selection in their experiment is risk dominance instead. In an experiment similar to ours, Rosenkranz and Weitzel (2012) compare the role of Nash tâtonnement stability with risk dominance and quantal response theory. Their findings provide partial support for all three theories, primarily because the rate of equilibrium play is very low, so that discrimination between the theories is difficult. Common to both these experiments, social preferences, and fairness considerations have never been given a chance to play a role as a coordination device. This is because most of their evidence stems from games on asymmetric network structure, where they find just the equilibrium that is predicted by all the equilibrium refinement concepts, including ours. Moreover, they either have a simple binary strategy space that rules out equal divisions by designing or implementing stranger’s designs that make it difficult for socially concerned players to coordinate their choices. Our design, in contrast, gives

⁴The exception here is Richefort (2018) who also studies a public goods game with strategic interaction. Nevertheless, his game yields a unique equilibrium whether or not players are socially concerned. Hence, also in his model, social preferences merely shift the equilibrium point.

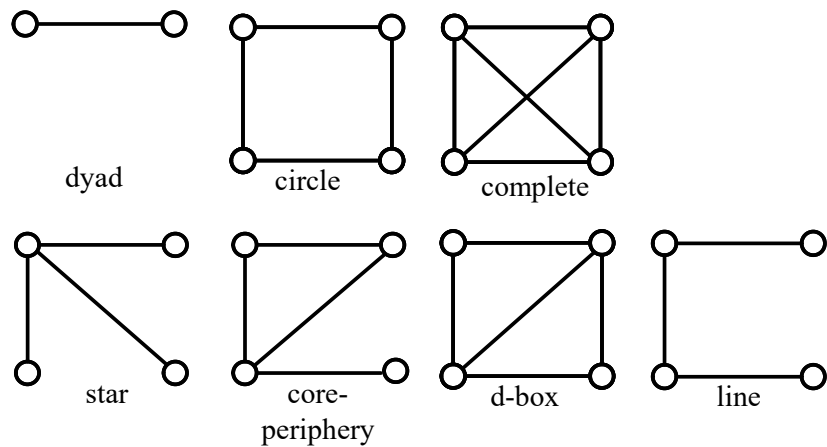


Figure 1: Networks in the experiment

social preferences a chance to play a role.

The only other experimental study of social preferences in network games that we are aware of is Zhang (2018). He compares the explanatory power of two preference types, altruism, and inequity aversion, in two network structures, star and circle, and concludes that altruism explains the data better. What we add is the structural test of a richer theory of social preferences on a larger set of networks.

Finally, our study is related to the large literature on social preferences. It is particularly close to a recently emerging group of studies that goes beyond the influence of pro-social behavior in standard public goods or bargaining games. These studies have already argued that it is probably too far-fetched to believe that social preferences would lead to radically different choices from payoff-maximizing behavior. Nevertheless, social preferences play an important role in these studies because they help to navigate unfamiliar social dilemmas (Binmore, 2005), are the foundation for social norms (Reuben and Riedl, 2013; Fehr and Schurtenberger, 2018), or create resistance to change (Eyster, Madarász, and Michailat, 2021). Closest to our study, Dufwenberg and Patel (2017) have shown that reciprocity concerns can facilitate coordination in a threshold-level public goods game with multiple equilibria. The arguments underlying their result are entirely different from ours. Moreover, while their theory speaks to public goods provisions in small communities, the application we have in mind is allocating scarce resources in a network of interdependent public goods.

3 Theory

3.1 Rules of the game

We study the role of social preferences in the Bramoullé and Kranton (2007) public goods game. The rules are as follows: n players are embedded in a fixed network g . Figure 1 illustrates the networks in our experiment.

All players simultaneously choose an investment that contributes to their

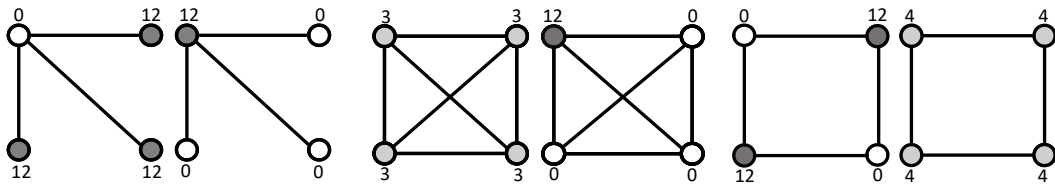


Figure 2: Nash equilibria on three networks

own local public good and to that of their direct neighbors in g .⁵ Let $e_{-i} = \{e_1, e_2, \dots, e_{i-1}, e_{i+1}, \dots, e_n\}$ denote the investments of all players except player i , and let $N_i = \{j \in N \setminus i : g_{ij} = 1\}$ denote the set of players in the neighborhood of i . The payoff of player i is given by

$$\pi_i(e_i, e_{-i}) = b\left(e_i + \sum_{j \in N_i} e_j\right) - ce_i, \quad (1)$$

where $b(\cdot)$ is the public-good benefit function, which is increasing and concave and which satisfies $b'(0) > c > b'(\infty)$ and $b''(x) \leq b''(y) < 0$ for $x \leq y$.

An important property of the Bramoullé and Kranton (2007) game is that investments are strategic substitutes. This means that there exists a strictly positive investment level e^* , such that all players aim to fill the gap between e^* and the investments in their neighborhood (given that the latter do not already exceed e^*).

Strategic substitutability furthermore implies that every network structure has multiple Nash equilibria that differ markedly in terms of the payoff distribution they induce. Figure 2 illustrates the equilibria for three of the networks in our experiment. On the star, for instance, there are two equilibria, one equilibrium where the center player invests e^* (given by $e^* = 12$ in our experiment) and the periphery players free ride; and another equilibrium where the periphery players invest each e^* and the center free rides. On the circle, there are three equilibria, one equilibrium where investments are distributed equally and two equilibria where every second player provides e^* so that every other player can free ride. In the complete network, there is even a continuum of equilibria because every investment profile is an equilibrium as long as the sum is equal to e^* . Moreover, also these equilibria differ markedly in terms of the payoff distribution they induce, with one equilibrium being even entirely symmetric. As will be shown, social preferences can lead to a fine-grained selection on these equilibrium sets, and the equilibria they select are empirically relevant.

⁵Examples of such partner-independent investments are organizing parties for friends, experimentation with new tools, or neighborhood beautification expenses, all vis-à-vis the time or effort a person spends on her personal “projects”.

3.2 A social preference function for network games

Social preferences are seen as the human tendency to take the payoffs of others into account when making a decision (e.g., Fehr and Schmidt, 1999). Yet, underneath this general tendency, there is much heterogeneity since people have been shown to differ widely in terms of their understanding of fairness (e.g., Bellemare, Kröger, and Van Soest, 2008; Falk, Becker, Dohmen, Enke, Huffman, and Sunde, 2018; Bruhin, Fehr, and Schunk, 2019).

The theoretical literature on social preferences has produced various meaningful utility functions that incorporate this empirical fact and that are suitable for two-player or n -player symmetric games (see Sobel, 2005, for a review). With the exception of the social network models mentioned in the literature review, games on more complex interaction structures have been left out of the perspective, however.⁶ Our preferred function is an n -player extension of the distributional preference function of Charness and Rabin (2002). It formulates a player's social preferences in the following way:⁷

$$U_i(e_i, e_{-i}) = \pi_i + \frac{1}{|R_i|} \sum_{j \in R_i} (\rho_i r_{ij} + \sigma_i s_{ij}) \pi_j, \quad (2)$$

where R_i denotes player i 's reference group, ρ_i and σ_i are i 's preference parameters with $1 > \rho_i \geq \sigma_i > -1$, and

$$\begin{aligned} r_{ij} &= 1 && \text{if } \pi_i \geq \pi_j \text{ and } r_{ij} = 0 \text{ otherwise,} \\ s_{ij} &= 1 && \text{if } \pi_i < \pi_j \text{ and } s_{ij} = 0 \text{ otherwise.} \end{aligned}$$

Utility is thus a linear combination of players' own material payoffs and a social preference component. The latter captures the (dis-)utility players derive from comparing their payoffs with those of other players. With whom a player compares is defined by her reference group R_i . This group might comprise any number of players. In a network context, it seems natural, however, that players compare with their direct neighbors who they can influence ($R_i = N_i$) or with everyone else in a network ($R_i = N \setminus i$).

Players thereby distinguish between peers who are equal or behind ($\pi_i \geq \pi_j$) and peers who are ahead ($\pi_i < \pi_j$). The parameters ρ_i and σ_i then govern the (dis-)utility from comparing with those behind and those ahead. In combination, they define various meaningful social preference types: Unconditional altruists ($\rho_i \geq \sigma_i > 0$), for instance, always put a positive

⁶These studies have developed their own interdependent utility functions. Our function nests several of them as special cases. In particular, Ghiglino and Goyal (2010) consider what we define as spitefulness, and Immorlica, Kranton, Manea, and Stoddard (2017) consider competitiveness. Broulès, Bramoullé, and Perez-Richet (2017) and Dufwenberg and Patel (2017), in contrast, develop models where players know each other well and accordingly, include each others' utilities rather than payoffs in their utility functions.

⁷We deviate from the original Charness and Rabin (2002) function in that the absolute level of another player's payoff enters utility function (2), rather than the relative payoff vis-à-vis the focal player. This modification circumvents a counter-intuitive prediction of the original function in the context of a continuous investment game like the Bramoullé and Kranton (2007) game. We address this issue in Appendix B.2.

weight on other players' payoffs. In the context of the Bramoullé and Kranton (2007) game, this means that altruists are always willing to maintain a public good in their neighborhood beyond the payoff-maximizing level e^* , regardless of whether their peers earn more or less. Also social-welfare types ($\rho_i > \sigma_i = 0$) like to maintain a public good beyond e^* , unless they earn less than everybody else. In that case, they behave like ordinary payoff maximizers aiming to fill the gap between their neighbors' investments and e^* . In the negative domain, spiteful types ($0 > \rho_i \geq \sigma_i$) always place a negative weight on other players' payoffs. They thus keep their investment below e^* to lower the payoffs of others. Competitive types ($0 = \rho_i > \sigma_i$), in contrast, refrain from these welfare-reducing deviations when their payoffs are weakly higher than everybody else. The two domains are connected by the inequity-averse types ($\rho_i > 0 > \sigma_i$) who are willing to increase their investments beyond e^* for players with a lower payoff and reduce their investments below e^* for players with a higher payoff.

Utility function (2) thus captures a wide range of empirically relevant preference types. Moreover, as we will see, it is simple enough to produce sharp predictions in the context of a continuous investment game.

3.3 General predictions

We now turn to our characterization of the equilibria with socially concerned players, henceforth the other regarding equilibria (ORE).

3.3.1 Rules of the modified game

In line with the setup in our experiment, and the empirical reality more broadly, we think of a game where each player has a different preference type. Specifically, suppose the parameters of utility function (2) are randomly determined for each player before the start of the game. The preference parameters of player i are summarized by her type $\tau_i \equiv (\rho_i, \sigma_i, R_i)$. Moreover, a combination of player types in the game is denoted by $\omega = (\tau_1, \tau_2, \dots, \tau_n)$. The set T_i then collects all potential types of player i , whereby we assume that T_i is a finite subset of all feasible types of function (2), and $\Omega = T_1 \times T_2 \dots \times T_n$ is the set of all potential type combinations.

Players learn their own types τ_i before the start of the game. In addition, they acquire some information about the preferences of the other players. Concretely, we make one of the following alternative assumptions:

- (A1) The player types τ_i are private information, but the player-specific type sets T_i are *common knowledge*.
- (A2) The player types are drawn from the common support T , but players have *complete information* about other players' types.

As becomes clear below, these assumptions on the information structure are important for social preferences to foster equilibrium selection. Nev-

ertheless, they are also reasonable to make. The common knowledge assumption, for example, is justified in any context where players know each other from prior encounters and have acquired at least some information about their preferences. All players need to have for our purposes is some vague impression about other players' preferences, such as player 1 is of a social-welfare or altruistic type, and player 2 is competitive or spiteful. The complete information assumption, in contrast, is stronger. Nevertheless, it can be justified in the context of our experimental games, where subjects are able to update their decisions for a certain time interval and have full information about the momentary investments of every other player. Because this is all a player needs to know to make her myopic best-response investment, the conditions in the experiment emulate a dynamic game with absorbing states that are identical to the Nash equilibria of a game that satisfies the complete information assumption.

3.3.2 Other-regarding equilibria

Let $f_i(\tau_i, e_{-i})$ denote the best-response investment of a type- τ_i player against the investments e_{-i} of every other player $j \neq i$. Under the incomplete information assumption, e_{-i} consists of one investment for each type $\tau_j \in T_j$ of every j . An ORE is then a (Bayesian) Nash equilibrium in pure strategies, where $e_{\tau_i} = f_i(\tau_i, e_{-i})$ for all players and all their types.

The following example illustrates the other-regarding best responses and the ORE for the simplest possible case.

Example 1. Suppose a two-player game with player 1 being a payoff maximizer and player 2 a social player, and suppose that both players have complete information about this. The best response of player 1 is given by

$$f_1(e_2) = \begin{cases} e^* - e_2 & \text{if } e_2 \leq e^* \\ 0 & \text{otherwise} \end{cases}.$$

The best response of player 2 depends, in contrast, on her payoff relative to player 1. Player 2's best-response correspondence thus has a kink in the middle:

$$f_2(\tau_2, e_1) = \begin{cases} e^*(\sigma_2) - e_1 & \text{if } e_1 \leq e' \\ e_1 & \text{if } e' \leq e_1 \leq \frac{1}{2}e^*(\rho_2) \\ e^*(\rho_2) - e_1 & \text{if } \frac{1}{2}e^*(\rho_2) < e_1 \leq e^*(\rho_2) \\ 0 & \text{otherwise} \end{cases},$$

where $e^*(\rho_i)$ and $e^*(\sigma_i)$ denote the total investments desired by a social player who earns more, respectively less, than the other player (defined in Lemma 1 below), and e' is a threshold value that satisfies $0 < e' \leq \frac{1}{2}e^*(\sigma_2)$.

As a result, player 2's social preferences constrain the equilibrium payoff

distribution. Moreover, depending on 2's preference type, the set of ORE might even be a strict subset of the payoff-maximizing equilibria:

$$(e_1, e_2) = \begin{cases} (e^*, 0) & \text{if } \tau_2 = \text{spiteful} \\ (\frac{1}{2}e^* \leq e_1 \leq e^*, e^* - e_1) & \text{if } \tau_2 = \text{competitive} \\ (\frac{1}{2}e^*, \frac{1}{2}e^*) & \text{if } \tau_2 = \text{inequity averse} \\ (\frac{1}{2}e^*, \frac{1}{2}e^*) \text{ or } (0 \leq e_1 \leq e', e^* - e_1) & \text{if } \tau_2 = \text{social welfare} \\ (\frac{1}{2}e^*, \frac{1}{2}e^*) \text{ or } (0, e^*(\sigma_2)) & \text{if } \tau_2 = \text{altruist} \end{cases} .$$

Point predictions for a more general network structure and arbitrary player types are more difficult to make. This is because the best-response investment of a social player can be significantly different from the investment that a payoff maximizer would make. Moreover, it depends on the player's relative standing vis-à-vis every single other player in her reference group.

Nevertheless, we can define several general properties that an other-regarding best response, and an ORE, must satisfy. First, the magnitudes of the parameters ρ_i and σ_i determine how far a player is willing to deviate from a payoff-maximizing best response. To measure this deviation, we define $\epsilon_i \in \mathbb{R}_+$ as the maximal absolute deviation that a type τ_i of player i would be willing to make from the investment that a payoff maximizer would do in her stead. Formally,

$$\epsilon_i \equiv \max \left\{ |f_i(\tau_i, e_{-i}) - f_i(e_{-i})| : \forall e_{-i} \in \mathbb{R}_+^{|\Omega-i|}, \forall g \in G \right\}. \quad (3)$$

In words, ϵ_i is a measure for the strength of a player's social preferences, and it is defined by the most "unfair" situation that a player may encounter in a game. The following result (proven in Appendix A.1) shows how the preference parameters can be mapped into a value for ϵ_i .

Lemma 1. *Suppose that players' utilities are defined by the social preference function (2) with parameters $\tau_i = (\rho_i, \sigma_i, R_i)$ and payoff function (1). A player's social preference strength, ϵ_i , is given by*

$$\begin{aligned} \text{altruist or social-welfare types} &: \epsilon_i^p = e^*(\rho_i) - e^* \\ \text{competitive or spiteful types} &: \epsilon_i^n = e^* - e^*(\sigma_i) \\ \text{inequity-averse types} &: \max\{\epsilon_i^p; \epsilon_i^n\}, \end{aligned}$$

and where $e^*(\rho_i) \equiv (b')^{-1}(\frac{c}{1+\rho_i})$ and $e^*(\sigma_i) \equiv (b')^{-1}(\frac{c}{1+\sigma_i})$.

Next, even though socially concerned players might deviate from a payoff-maximizing best response, they do not deviate in an arbitrary way. As they ultimately strive for a certain payoff ordering, the investment profiles that qualify as an ORE are constrained in a systematic way. To see how they are constrained, let us look at the following example.

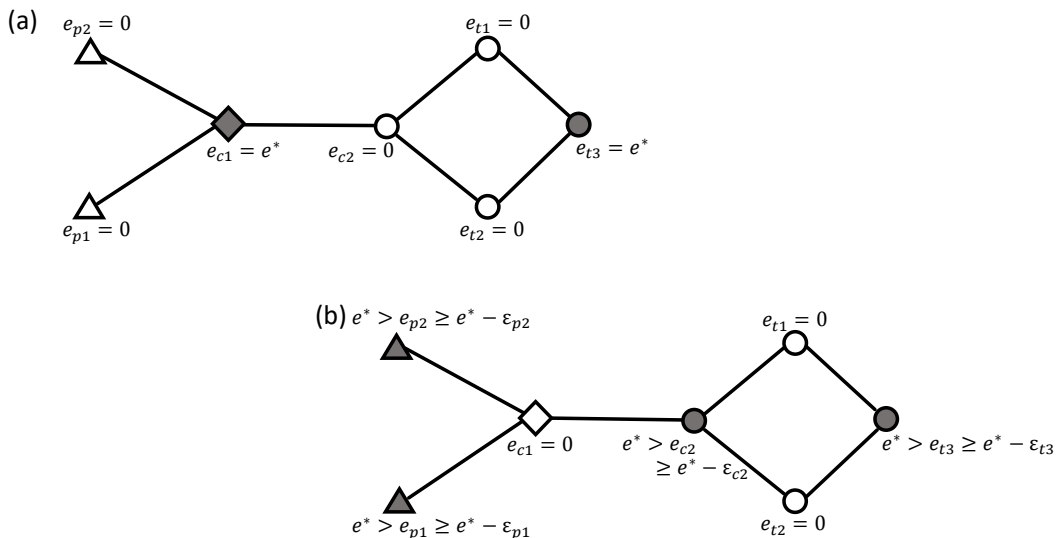


Figure 3: Payoff-maximizing and refined other-regarding equilibria

NOTES: Panel (a) shows one of the four specialized payoff-maximizing equilibria that is *not* a refined ORE. Panel (b) shows a *refined ORE* that coincides with one of the other three specialized payoff-maximizing equilibrium profiles when $\epsilon \equiv \max\{\epsilon_i \mid i \in N\} \rightarrow 0$. The other specialized payoff-maximizing equilibrium that can be supported in a refined ORE is $e^* > e_i \geq e^* - \epsilon_i$ for $i \in \{p1, p2, t1, t2\}$ and $e_{c1} = e_{c2} = e_{t3} = 0$. The triangular and diamond nodes indicate players in nested neighborhoods.

Example 2. Consider the investment profile in Figure 3 Panel (a). Suppose that all players are inequity-averse and suppose they only compare with their direct neighbors, both of which is known by all players (i.e., assumption (A2) applies). The profile in Panel (a) cannot be maintained in an ORE in this case, despite being a payoff-maximizing equilibrium. This is because players $c1$ and $t3$ would want to reduce their investments below e^* as they are the only ones who contribute in their neighborhood and therefore feel exploited. At the same time, all other players would want to increase their investments (from zero) because they feel guilty.

To construct an ORE, we must therefore seek for an investment profile where player $c1$ earns more than at least one of her neighbors, say player $p1$ because player $c1$'s envy of players $p2$ and $c2$ may then be balanced out against her guilt towards $p1$. Such a profile is displayed in Panel (b). Here, player $c1$ free rides entirely on the investments of her neighbors, who each make a positive contribution. To make this an ORE, we additionally require, however, that player $c1$ and the other two free-riding players $t1$ and $t2$ see no reason to make an investment despite their guilt. Put differently, we require that players $t1$ and $t2$ receive more from their neighbors than the highest investment to which they are willing to contribute. This is certainly the case when

$$2(e^* - \epsilon) \geq e^* + \epsilon \Leftrightarrow \epsilon \leq 0.33 e^*,$$

where $\epsilon \equiv \max\{\epsilon_i \mid i \in N\}$ denotes the maximal preference strength in the player group. The corresponding condition for player $c1$ is $3(e^* - \epsilon) \geq e^* + \epsilon \Leftrightarrow \epsilon \leq 0.5 e^*$.

A striking feature of the ORE in Panel (b) is that the central players $c1$

and $c2$ are treated differently. While $c1$ earns more than everybody else, $c2$ is one of the players who earns the least. The reason lies in their different positions in the larger network. Specifically, unlike player $c2$, player $c1$ *nests* the neighborhoods of her neighbors $p1$ and $p2$ who thus do not receive any investment which $c1$ does not have access to. Combined with their aversion to inequity, $p1$ and $p2$ thus cannot earn more than player $c1$ because their feelings of guilt would make them increase their investments e_{p1} and e_{p2} beyond what a payoff-maximizer would invest so that together with $c1$'s investment we get $e_{pi} + e_{c1} > e^*$. At the same time, player $c1$ cannot earn weakly less than everybody else, because she would reduce her investment below what a payoff-maximizer would invest in her stead so that $e_{c1} + \sum_{j \in N_{c1}} e_j < e^*$. These two conditions contradict each other, however.

The same cannot be said about player $c2$ and her neighbors in Panel (b). All three of them have access to at least one other player, who $c2$ does not have access to and who contributes to their public good. And, because the total investments that $c2$'s neighbors receive are beyond their personally desired level of the public good when $\epsilon \leq 0.33 e^*$, they are not willing to make the extra investment that would be needed for a more equitable outcome. In an ORE, $c2$ can thus be one of the players who earn the least.

Equilibrium selection through social preferences is tied to another additional condition, however, next to the network's nestedness. To see which, consider the profile in Panel (a) again.

Example 3. Suppose that, instead of all players being inequity-averse, players $c1$ and $t3$ are of a social-welfare type while all other players are competitive or spiteful. The profile in Panel (a) can then be maintained in an ORE—in addition to the profile in Panel (b)—because $c1$'s and $t3$'s neighbors do not feel guilty any longer (maintain $e_i = 0$), while $c1$ and $t3$ look after themselves (play $e_j = e^*$).

Why can the profile in Panel (a) been ruled out as an ORE in Example 2 but not in Example 3? The reason is that in Example 2, players' social preferences "fit" the network position they are in, whereas this is not the case in Example 3. Broadly speaking, equilibrium selection through social preferences is facilitated by competitive or spiteful types in the nesting positions of a network because these types are determined to undo any payoff difference in their disadvantage if there is need to. Equilibrium selection is, on the other hand, fostered by social-welfare types or altruists in the nested positions of a network because these types are willing to undo any payoff disadvantages for their neighbors. Inequity-averse types, finally, support equilibrium selection in any network position because these types are willing to undo their own payoff disadvantages as well as those of their neighbors. Generalizing from here, we say that two players have compatible social preferences when their types satisfy the following condition:

Definition 1. Consider two neighbors i and j in a network such that i nests the neighborhood of j , that is, it holds $N_j \cup j \subseteq N_i \cup i$. We say that their social preferences are compatible if $\tau_i \in T_i^*$ and $\tau_j \in T_j^*$ such that

$$T_i^* = \{\text{competitive, spite, inequity averse}\} \quad \text{and} \quad T_j^* = T \setminus \{\text{spite}\} \\ \text{OR} \tag{4}$$

$$T_i^* = T \setminus \{\text{altruist}\} \quad \text{and} \quad T_j^* = \{\text{social welfare, altruist, inequity averse}\}.$$

Based on this definition, we say that an other-regarding equilibrium is a *refined ORE* when at least *some* pairs of players in nested neighborhoods have compatible social preferences.

The following result synthesizes the insights from Examples 2 and 3 into a general property of a refined ORE. Let $\pi_i(\omega)$ denote the equilibrium payoff of the type τ_i of player i associated with the type combination ω . Then,

Proposition 1. Consider two players i and j in a network such that i nests the neighborhood of j . Moreover, suppose that either information assumption (A1) or (A2) applies. In a refined ORE, there exists at least one type combination $\omega \in \Omega^* = T_1 \times \dots \times T_i^* \times \dots \times T_j^* \times \dots \times T_n$ such that the type of player i (j) earns weakly more (strictly less) than at least one other player k (l) in the player's reference group. That is,

$$\pi_i(\omega) \geq \min_{k \in R_i} \{\pi_k(\omega)\} \quad \text{OR} \quad \pi_j(\omega) < \max_{l \in R_j} \{\pi_l(\omega)\} \tag{5}$$

for at least one $\omega \in \Omega^*$.

The intuition behind Proposition 1 (proven in Appendix A.2) extends immediately from Examples 2 and 3. Suppose that payoff condition (5) is violated, then this immediately implies that $\pi_i(\omega) < \pi_j(\omega)$. Because player j would feel obliged to “help” player j out in this case, j would choose an investment above the payoff-maximizing best-response level. At the same time, since player i would feel “exploited”, she would reduce her investment below the payoff optimal level. Combined with the fact that i nests the neighborhood of j , this leads to a contradiction to best-response behavior, however, because i has access to more investments than j while i invests relative little. Thus, in a refined ORE, the payoffs of i and j must be ranked according to (5).

Note that the assumption that *all* types of players i and j are drawn from the restricted sets T_i^* and T_j^* is crucial in this argument. Otherwise, there could be a type of player j who is *not* willing to help the focal type of player i so that i could not afford to reduce her investment. The assumption of either common knowledge (A1) or complete information (A2) about the types of other players is also important. Otherwise, a player j of the correct type $\tau_j \in T_j^*$ might still mistakenly believe that player i is not needy or player i might believe that j is not willing to help, etc. Equilibrium selection through

social preferences thus requires at least some common understanding between players.

3.4 Experimental predictions

In the following, we fully characterize the sets of refined ORE for the seven networks in our experiment. For ease of exposition, we thereby focus on the predictions for the complete information assumption (A2) as these are more easily linked to the setup of our experiment. The proofs of all statements can be found in Appendix A.3. Table 6 in the same appendix summarizes the predictions, again.

3.4.1 Dyad and complete network

Remember that any investment profile can be supported in a payoff-maximizing equilibrium on the dyad or the complete network as long as the sum of investments is equal to e^* . Social preferences expand this equilibrium set, in a first instance, because deviations up to $\pm\epsilon$ from e^* are supported in an ORE as well.

When the social preferences of all players are compatible, however, then these players will coordinate on a very fine-grained subset of the equilibria. In particular, suppose the preference requirements of Appendix A.3 are met by every player on the dyad or complete network. Then, all players must invest exactly the same,

$$e_i = e_j = e, \text{ where } e \in \left[\frac{e^* \pm \epsilon}{n} \right]. \quad (6)$$

The intuition is as follows. Suppose that, contrary to (6), not all investments are equal. The fact that players' neighborhoods are mutually nested means that the players with the highest investment earn weakly less than everybody else and the players with the lowest investment earn weakly more. At least one player would thus feel insulted in her understanding of fairness and adjust her investment up- or downward. Such adjustments can only be avoided when all players invest just the same.

3.4.2 Star, core periphery, and d-box

Two markedly different investment profiles can be supported in an equilibrium on the star when all players are payoff maximizers: a periphery-sponsorship profile where the center player free rides on the peripheral players who each contribute e^* ; or a center-sponsorship profile where the peripheral players free ride on the center's investment of e^* . The equivalent

ORE on the star network satisfy

$$\begin{aligned} \text{(periphery spon.)} \quad & e_i = 0 \quad \text{and} \quad e^* - \epsilon \leq e_j \leq e^* + \epsilon \\ \text{(center spon.)} \quad & 12 - \frac{7\epsilon}{3} \leq e_i \leq 12 + \epsilon \quad \text{and} \quad \sum_{j \in P} e_j \leq 4\epsilon, \end{aligned}$$

where i denotes the center player and $j \in P$ the peripheral players. The ORE on the core-periphery network and the d-box have a very similar structure (see Table 6 in Appendix A.3).

In contrast, when players have compatible social preferences, a pure center-sponsorship equilibrium with $e_j = 0$ for all $j \in P$ can be ruled out. This is because the center player(s) must earn weakly more than at least one non-center player in a refined ORE. Formally,

$$\pi_i(e) \geq \min_{j \in N \setminus C} \{\pi_j(e)\} \quad \text{for all } i \in C. \quad (7)$$

Suppose moreover that all players have limited preference strengths in addition (see Appendix A.3 for the bounds). Then, the set of ORE can be refined even further because the above payoff ordering can only be guaranteed for when the public good is entirely sponsored by the non-center players.

3.4.3 Line

With payoff-maximizing players on the line network, every investment profile is a Nash equilibrium as long as it satisfies for middle player mi and for end player ei :

$$e_{ei} = e^*, \quad e_{mi} = 0, \quad e_{m-i} + e_{e-i} = e^* \quad \text{for } i \in \{1, 2\}.$$

Expanding on this, the set of ORE also consists of end-sponsored and distributed public goods. A refined ORE, in contrast, must be an end-sponsored public good with

$$\pi_{mi}(e) \geq \pi_{ei}(e) \quad \text{for } i \in \{1, 2\}. \quad (8)$$

Thus, social preferences also select among the equilibria on the line when they match the network positions which the players occupy. Yet, they do so less effectively than on the star, core periphery, or d-box because the payoff ordering they imply only applies to the end players and their direct neighbors in the line middle.

3.4.4 Circle

The absence of nested neighborhoods puts an end to the equilibrium selection property of social preferences on the circle. All that can be said is

summarized in the definition of ϵ in expression (3): The set of ORE is wider than the payoff-maximizing equilibrium set and collapses with it if $\epsilon \rightarrow 0$. In particular, when ϵ is small, other-regarding players coordinate on either a (near) *distributed* investment profile or a (near) *specialized* profile where every second player free rides on the investments of her neighbors.

Maybe surprisingly, a specialized profile can even be supported on the circle when all players are social-welfare concerned or inequity-averse. This follows from what we said about the role of nestedness. Even though the contributing players might feel exploited, they maintain their investments for the sake of their own payoffs. The free riders, therefore, receive a total contribution beyond their personal desired level of the public good and consequently see no reason to bear the extra cost of a more equal outcome.

3.4.5 Network ranking

So far, we have seen that adding social preferences to the Bramoullé and Kranton (2007) game allows us to exclude several of the investment profiles that were Nash equilibria in the standard game where players are payoff maximizers. Moreover, when the social preferences of our players are small ($\epsilon \rightarrow 0$), we even obtain an ORE set that is a strict subset of the payoff-maximizing equilibria for many of the networks in Figure 1.

Our theory does predict more, however. It suggests marked differences between the networks of Figure 1 in terms of how likely a group of players will coordinate on a refined ORE. A first observation is that for a randomly drawn type combination ω from the entire set $\Omega = T^n$, the likelihood that ω yields a combination of compatible preference types declines, *ceteris paribus*, with the size of a player group. This has some immediate implications for the dyad and the complete network because it means that

$$P(\omega \in \Omega^{dyad}) \geq P(\omega \in \Omega^{comp}), \quad (9)$$

where Ω^{dyad} and Ω^{comp} denote the sets of compatible preference combinations for these networks. Hence, we expect to see more equal-split profiles on the dyad.⁸ That type of coordination problem is well known in the literature (e.g., Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000). Coordination is harder to achieve in larger groups because it is more difficult to get a sufficient number of players together who share a common understanding about which equilibrium to play.⁹

In a network context, the likelihood of yielding a compatible preference

⁸For this, we implicitly make the plausible assumption that players coordinate on a random profile from the set of profiles compatible with their preferences, that is, the wider set of ORE when players have incompatible preferences and the narrower set of refined ORE when their preferences are compatible.

⁹In the theory of Fehr and Schmidt (1999), for example, the number of players adversely affects the likelihood of cooperation in a public goods game with punishment options because the likelihood that a sufficient number of conditional cooperators is present is smaller in larger groups.

combination also depends on the precise structure of a network. We have already seen above that in the absence of nested neighborhoods, such as in the circle, a network is prone to produce multiple equilibria even when all players have homogeneous preferences. Thus, at least some degree of nestedness is a necessary prerequisite for a common understanding about which equilibrium to play.

But even among the nested networks of Figure 1, our theory predicts some marked differences. In particular, there is some asymmetry with regard to the ideal number of central players (n_c) who nest other players' neighborhoods, and the ideal number of peripheral players (n_p) whose neighborhoods are nested. The larger n_c (e.g., comparing the star and the d-box), the more likely it is that there is one central player who is of a social-welfare or altruistic type so that this player is willing to sponsor the public good when no one else in her neighborhood does. The larger n_c , therefore, the smaller the likelihood of a refined ORE as defined in (7) or (8). The number of peripheral players has the opposite effect. The larger n_p , the more likely it is that at least one of them is willing to help out the center players because this peripheral player is of a social-welfare or altruistic type. The larger n_p , therefore, the higher the likelihood of a refined ORE. Applied to our networks, we thus get

$$\begin{aligned} P(\omega \in \Omega^{star}) &\geq \left\{ P(\omega \in \Omega^{core}), P(\omega \in \Omega^{dbox}) \right\} \\ &\geq P(\omega \in \Omega^{line}). \end{aligned} \quad (10)$$

This means that coordination on a refined ORE is easiest on the star, followed by the core periphery and d-box, and followed by the line. Furthermore, coordination is easier on an asymmetric than on a symmetric nested network:

$$P(\omega \in \Omega^{core}) \geq P(\omega \in \Omega^{comp}). \quad (11)$$

Altogether, our theory leads to the following testable predictions:

Hypothesis 1: In the networks of Figure 1, except the circle, a group of players with compatible social preferences is more likely to coordinate on a refined ORE than a group without compatible preferences.

Hypothesis 2: The likelihood of a refined ORE depends on the network structure. In particular, the networks of Figure 1 can be ranked according to the conditions in (9)–(11).

Finally, for the circle network, we expect that even if all the compatibility criteria of Definition 1 are met by a group of players, the group does nevertheless *not* coordinate more likely on either a specialized or a distributed profile than a group who does not meet the criteria.

4 Experiment

We tested our hypotheses in an experimental version of the Bramoullé and Kranton (2007) game. In particular, we modified the original static game because earlier experiments made it clear that subjects find it very hard to coordinate their choices in this game. Coordination was particularly difficult in experiments that adopted the original large strategy space (e.g., Rosenkranz and Weitzel, 2012). As at least some equilibrium play is essential for our theory testing, however, we opted for a dynamic extension of the game that nevertheless retains some key properties of the original versions.

Specifically, following Callander and Plott (2005), Berninghaus, Ehrhart, and Ott (2006), and Goyal, Rosenkranz, Weitzel, and Buskens (2017), every game lasted between 30 and 90 seconds in our experiment. The final decision moment, t^{max} , was randomly determined by a draw from a uniform distribution on the support $[30, 90]$. Starting from a situation of zero investments, subjects could continuously update their investments, choosing from the entire set of positive integer values. Full information about the momentary investments of all other players was continuously provided and updated five times per second. The momentary payoffs were, moreover, indicated by the size of each player's node on the screen (see screenshot in Appendix C.2). Nevertheless, the actual payoff was solely determined by the momentary investments at the random round end. Payoffs were thereby calculated as follows

$$\pi_i = \begin{cases} (e_i + \sum_{j \in N_i} e_j)(29 - e_i - \sum_{j \in N_i} e_j) - 5e_i & \text{if } e_i + \sum_{j \in N_i} e_j \leq 14 \\ 196 + e_i + \sum_{j \in N_i} e_j - 5e_i & \text{otherwise} \end{cases}.$$

As we will see below, coordination was greatly facilitated by these design choices essentially because subjects did not need to formulate beliefs about other players' payoffs and investments. At the same time, the implemented random stopping rule avoided last round effects.

4.1 Experimental procedure

We administered our experiment at the Experimental Laboratory for Sociology and Economics (ELSE) at Utrecht University, the Netherlands. The experiment was programmed in z-tree 3.0 (Fischbacher, 2007) and subjects were recruited via ORSEE (Greiner, 2015). A total of 120 students participated in eight sessions, with 12–20 students each session. No subject could attend more than one session. Moreover, the order of games was randomly varied between sessions.

The average subject's age was 22, 67% were female, and 72% were of Dutch nationality. Subjects played each of the seven networks of Figure 1 in a total of 35 games, one trial game and four payoff-relevant games per

network. In each game, subjects were assigned to a random group of two or four players and a random network position. The entire experiment thus consisted of 960 payoff-relevant network games: 120 games per four-player network and 240 games for the dyad. A typical session lasted 80 minutes, and subjects earned 11.82 euros on average, including a 3 euro show-up fee.

4.2 Social preference elicitation

Key to our testing of Hypotheses 1 is that we have an estimate for the social preference parameters of our subjects. We estimated the parameters directly from their behavior in the network games.

In doing so, we confined the set of network games for our estimations in two ways. First, we ensured that we did *not* use the same games for our estimations that we used to test our theory on. In particular, every time we tested the conformity of a group's behavior with our predictions, we estimated the members' social preferences from their investments in another set of network games (with different players). Second, we restricted the set of games to ensure a balanced set of network positions for each subject. This is because some positions, for instance the periphery positions of the star, are due to the random assignment to network positions over-represented in a subject's set of games compared to, for example, the star center position. Based on theory, we expect, however, that these two positions confronted subjects with different decision situations and therefore triggered a different social comparison concern.¹⁰ We thus categorized the network positions into three classes:

- center positions (of the star, line, core periphery, and d-box)
- periphery positions (of the same networks), and
- symmetric positions (of the circle, core duo, complete network, and dyad).

We then estimated each subject's social preferences on an equal number of games from each class. Otherwise, we used as many payoff-relevant decision moments ($t \in [30, t^{max}]$) and as many network games as possible.

Obviously, we made several choices. Therefore, to check the robustness of our findings, we also elicited subject preferences in several alternative ways. Appendices B.2 and B.3 summarize our findings on these tests.

¹⁰This was confirmed in a pre-test where we estimated the subject-average $(\hat{\rho}_i, \hat{\sigma}_i)$ -pair per network position and found that the average greatly differed by position.

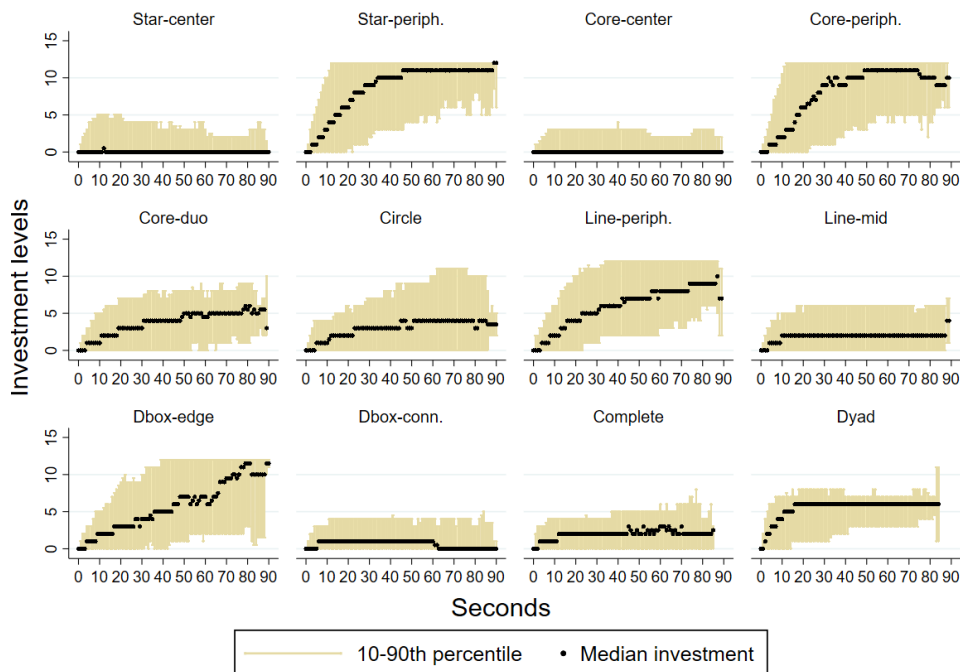


Figure 4: Investments by network position over time

5 Results

5.1 Descriptive findings

We first give an overview about the behavior in our experimental games before we test our hypotheses. If not stated otherwise, all results refer to the final investments at the randomly determined game ends.

Position-level findings: Figure 4 plots for each network position the evolution of the median investment and the 10–90 percentile over time. Clearly, with the exception of the d-box edge position, the medians converge to some steady state values. At the same time, the 10–90 percentile ranges shrank in all network positions but the circle, with a 90th percentile that is always lower than the payoff-maximizing public good level of $e^* = 12$ across all network positions. Thus, the evolution of behavior in our games is reminiscent of some best-response dynamic that converges to a static equilibrium.¹¹

In support of this, Figure 5 plots the distribution of deviations from a payoff-maximizing best response across all network positions and games. Clearly, the predominant choice is a pure payoff-maximizing best response: 40% of investments. And, when subjects deviated from it, they typically did not deviate by much: 74% of choices are no more than ± 2 units away from a payoff-maximizing best response, whereby subjects typically deviated in the downward direction.

For more detail, Figure 6 plots the distributions of final investments per network position. In line with our static ORE predictions for the dyad and

¹¹The disturbance in this pattern after 70 seconds in some network positions is mainly due to the fact that many games ended before that time.

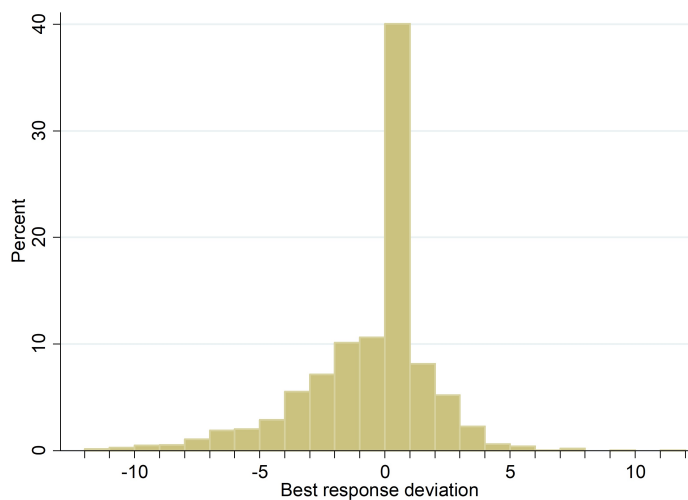


Figure 5: Deviations from payoff-maximizing response

NOTES: $N = 3,359$ final investment decisions in all games are compared to a payoff-maximizing best response. One value [+24] dropped for better display.

the complete network, the unique distributional modes are at six and three units respectively, which is consistent with the predicted equal-split equilibrium. Moreover, subjects found it considerably harder to coordinate their choices on the complete network, which is in line with the aggravated coordination problem predicted in Hypothesis 2.

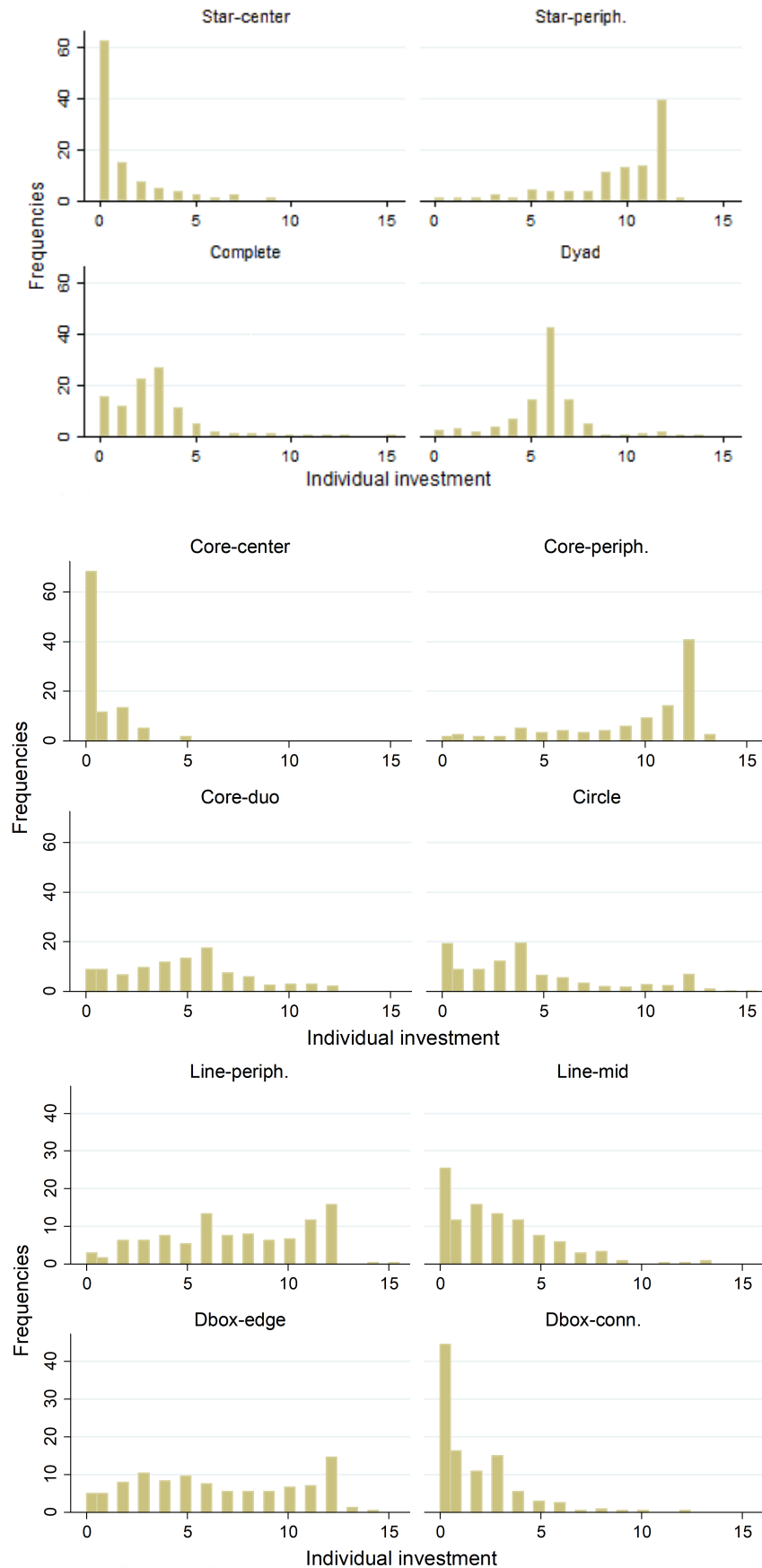
Turning to the asymmetric networks, consistent with our predictions for the center positions of the star, core periphery, d-box, and—to a lesser extent—the line, the most frequent choice is the zero contribution. Subjects in the peripheral positions of these networks, in contrast, oftentimes invest twelve units. Together, this suggests the predicted periphery-sponsorship equilibrium, whereby the pattern is more pronounced on the star and the core periphery network as predicted by Hypothesis 2.

For the circle network, Figure 6 shows a very dispersed investment pattern. Nevertheless, what we see does not go against our theory, as we predicted that players might either coordinate on one of the two possible specialized equilibria or a distributed equilibrium. Quite on the contrary, if we consider deviations of up to one (two) unit(s) from a payoff-maximizing best response as consistent with an other-regarding response, 78% (95%) of the investments in Figure 6 can be rationalized.

In sum, the first look at our findings by and large supports our theoretical predictions. Nevertheless, because the investments of all players need to “fit” in equilibrium, we now turn to the group-level behavior.

Group-level findings: Table 1 presents the shares of investment profiles consistent with a (refined) ORE for each network. We thereby distinguish between three degrees of deviation from a pure payoff-maximizing equilibrium: ORE with zero ($\chi = 0$), two units ($\chi < 3$), and any units (any χ) of deviation from a payoff-maximizing best response by at least one player.

Figure 6: Investments by network position



NOTES: Observations in star center, core center, core periphery, line middle, and line periphery: 120; core duo, d-box center, and d-box edge: 240; star periphery: 360; dyad, complete, and circle: 480. One value on the dyad [29] dropped for better display.

The exact criteria can be found in Table 6 in Appendix A.3.¹²

¹²The critical value $\chi < 3$ is chosen because a deviation of up to two units is the maximum deviation for which a periphery-sponsored public good is the unique refined ORE

Table 1: Frequencies of other-regarding equilibria

Network	Equilibrium type	Deviation from payoff-maximizing equilibrium		
		zero ($\chi = 0$)	moderate ($\chi < 3$)	any (any χ)
Dyad	egalitarian (rfd)	32.1%	45.8%	49.2%
	other	8.8%	33.0%	50.8%
Complete	egalitarian (rfd)	0.8%	0.8%	0.8%
	other	20.8%	62.5%	99.2%
Star	per-spon. (rfd)	15.8%	33.3%	62.5%
	cent-sp. with $\pi_c \geq \pi_j$ (rfd)	—	—	36.6%
	cent-sp. other	0%	0.8%	0.8%
Circle	specialized	7.5%	16.6%	29.2%
	distributed	3.3%	27.5%	55.0%
Core	per-spon. (rfd)	17.5%	43.3%	68.3%
	cent-sp. with $\pi_c \geq \pi_j$ (rfd)	—	—	31.7%
	cent-sp. other	0%	0%	0%
D-box	per-spon (rfd)	8.3%	15.0%	25.8%
	cent-sp. with $\pi_c \geq \pi_j$ (rfd)	—	0.8%	64.2%
	cent-sp. other	0%	3.3%	10.0%
Line	end-spon. (rfd)	0.8%	10.0%	46.7%
	distr. with $\pi_m \geq \pi_e$ (rfd)	8.3%	26.7%	39.2%
	distr. other	0.8%	0.8%	14.1%

NOTES: Percentages of investment profiles consistent with an other-regarding equilibrium (ORE) at the random ends of the 960 network games. 240 observations for dyad, 120 for all other networks. Refined ORE are indicated with "(rfd)".

A first observation is that the number of groups converging on a payoff-maximizing equilibrium (ORE with $\chi = 0$) is remarkably high.¹³ Not surprisingly, the number of ORE is even higher when investments in the neighborhood around a payoff-maximizing equilibrium point are rationalized by subjects' social preferences. Interestingly, however, a small expansion of the range of feasible profiles is already enough to capture a significant share of observed investments. For example, an inclusion of a deviation of ± 2 units ($\chi < 3$) adds meaning to the frequently observed downward deviations in the periphery positions of the star, core periphery, line, and d-box (see Figure 6). As a result, the shares of profiles consistent with an ORE more than doubles on these networks.

Turning to the evidence regarding the refined ORE, we first look at Column 3 (with $\chi = 0$):

- (i) On the asymmetric networks (star, core periphery, d-box, and line), we observe almost exclusively periphery-sponsored public goods. The only exception is the 0.8% of groups on the line network (i.e., exactly one group) who converge to a distributed equilibrium.

in all the asymmetric networks, except the d-box for which the critical value is already at $\chi < 2$ (see Table 6).

¹³Compared to Rosenkranz and Weitzel (2012), for example, who study the same games with a non-continuous-time design, the number of groups converging on a payoff-maximizing equilibrium increases by a factor of 3.4 in the star network (smallest increase) to 27 in the circle (largest increase).

- (ii) On the dyad, a large majority of groups (32.1%) splits $e^* = 12$ equally. In the complete network, this is the case for only 0.8% of groups (i.e., exactly one group), which is however not entirely surprising given the predicted coordination problems.
- (iii) On the circle, 7.5% of groups coordinated on a specialized equilibrium with alternating investments of zero and twelve units. Another 3.3% of groups coordinated on an equal-split equilibrium. Thus, as expected, both types of payoff-maximizing equilibria can be found.

In sum, the numbers in Column 3 largely support our theory because if an investment profile is consistent with a payoff-maximizing equilibrium it is almost always a refined ORE. Similar pictures emerge when we look at the wider sets of ORE in Columns 4 and 5. On the star, core, d-box, and line, the vast majority of groups converge to either a pure periphery-sponsored public good or a partially center-sponsored public good where the center player earns more than at least one periphery player. On the dyad, almost half the groups choose an equal-split equilibrium. Finally, on the circle, both the shares of “nearly” specialized and “nearly” distributed profiles increase significantly when we look at the wider sets of ORE. So, again, the only exception is the complete network where the share of equal-split equilibria remains at 0.8%.

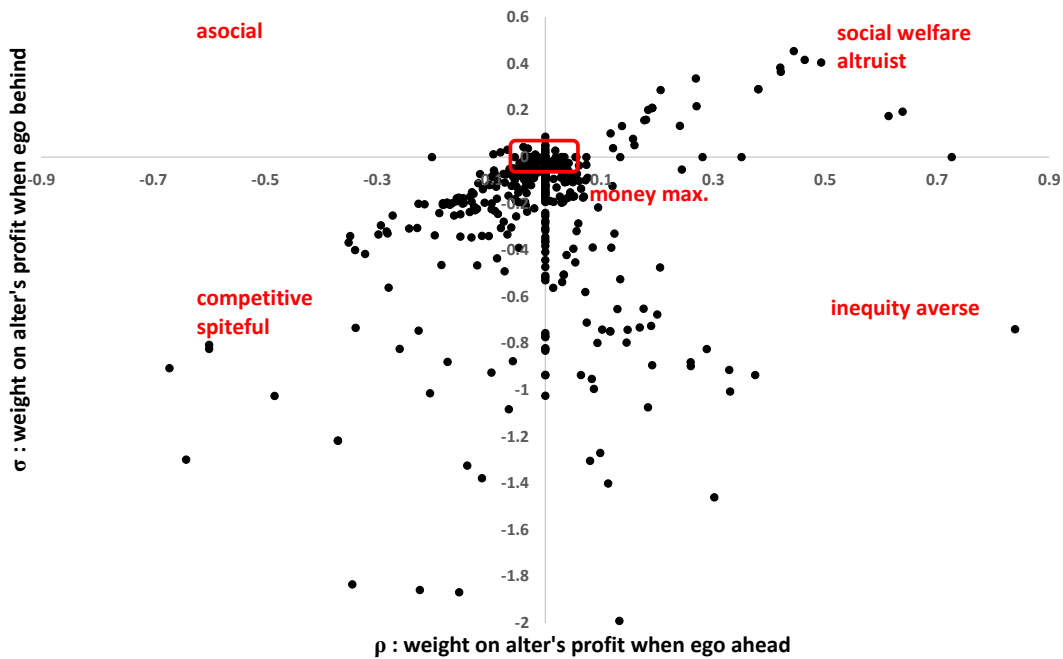
To put these findings into perspective, Appendix B.1 compares the numbers in Table 1 with the predictions and hit rates (Selten, 1991) of several alternative equilibrium refinement concepts, notably efficiency, Nash tâtonnement stability, and quantal response theory. To sum up the findings, our social preference theory predicts at least as well as the best alternative refinement concept on all network structures. Its specific power is that it selects the “natural” equilibria on the dyad and all the asymmetric networks (star, core periphery, d-box, line) and, at the same time, it does not rule out the co-existence of multiple, empirically relevant equilibria on the circle.

5.2 Hypothesis 1: preference compatibility

We have seen above that the behavior in our experiment is much in line with our theory. Nevertheless, we cannot yet rule out the possibility that our subjects chose their investments for reasons other than subjects’ social comparison concerns. Much of what we observed may, for instance, also be explainable with one of the refinement concepts mentioned above.

To remove all doubts, we test here a unique prediction of our theory that discriminates it from all conceivable alternative explanations: The reason why one subject group coordinates on one of the frequently observed refined ORE, and another group does not, is that the former has a set of compatible social preferences.

Figure 7: Social preference estimates



NOTES: Estimated $(\hat{\sigma}_i, \hat{\rho}_i)$ from 480 subject- and game-specific conditional logit estimations of function (2) assuming that subjects only include their direct neighbors in their reference group, i.e., $R_i = N_i$. See Section 4.2 for procedural details. Ten pairs (with $\hat{\sigma}_i < -2$) are dropped for better display.

Social preference estimates: To test this hypothesis, we first need an estimate for the social preference parameters of our subjects. We performed the subject- and game-specific conditional logit estimations for this purpose, which we explained in Section 4.2. Concretely, we assumed that a subject chose an investment e_i at time t of a network game so as to maximize utility function (2) augmented by a random utility component and responding to the observed investments of the other players at $t - 1$. We then estimated the $(\hat{\rho}_i, \hat{\sigma}_i)$ -pair that maximizes the conditional likelihood that this subject chose the actual investment pattern across all her games rather than any feasible alternative investment pattern. For practical reasons, we limited the alternative investments to $\{0, 1, 2, \dots, 15\}$ in our estimations. Moreover, for our main specification, we assumed that subjects only include their direct neighbors in their reference group.

Figure 7 visualizes the estimated $(\hat{\rho}_i, \hat{\sigma}_i)$ -pairs. Table 9 in Appendix B.2 categorizes them into the social preference types identified by function (2). Table 10 in the same appendix summarizes the preference estimates from several alternative utility functions and an alternative subset of data, including a specification where subjects compare their payoffs with everyone in the network. Much in line with the findings from earlier experiments (e.g., Falk, Becker, Dohmen, Enke, Huffman, and Sunde, 2018; Bruhin, Fehr, and Schunk, 2019), both our preferred estimates in Figure 7 as well as the alternative estimates in Appendix B.2 suggest considerable subject heterogeneity in social preferences. In particular, there are sizable numbers of subjects who showed a concern for each of the preference types in (2).

In a next step, we could thus classify each of the 840 subject groups in

Table 2: Groups with compatible preferences

Pref. strength	dyad	star	core	d-box	line	complete
any $\hat{\epsilon}$	22.9%	70.8%	73.3%	30.0%	10.8%	3.3%
$\hat{\epsilon} < 3$	17.1%	40.0%	45.8%	16.7%	7.5%	2.5%
$\hat{\epsilon} < 1$	14.2%	25.0%	26.7%	6.7%	5.0%	2.5%
No. of groups	240	120	120	120	120	120

NOTES: Categorization of 840 subject groups (all but the 120 groups playing the circle) according to whether their members meet the preference-compatibility requirements for a given network or not. Groups are additionally classified by the maximum preference strength of their members. The detailed requirements can be found in the network characterizations of Appendix A.3 and in Lemma 1.

our experiment (120 per 4-player network excluding the circle, 240 for the dyad) with regard to whether all the group members have a combination of compatible preferences or not. The results of this classification can be found in Table 2. Clearly, the table suggests that there is a sizable number of groups for every network structure who meet the preference requirements of Section 3.4 and another sizable number of groups who do not meet these requirements. We can thus turn to our main question whether groups who do meet the requirements play a refined ORE more often.

Hypothesis test: Table 3 reports the results of two multinomial logit models. The dependent variable in these models is a multinomial variable that classifies the investment profiles chosen by a subject group at the final decision moment of a game into six different possible outcome types. Outcomes (1)–(3) are the network-dependent refined ORE as predicted by our theory, outcomes (4)–(6) are the remaining non-refined ORE. The two outcome classes are further subdivided into how much a group’s investments deviated from a pure payoff-maximizing equilibrium. Together, these six outcomes capture all feasible investment profiles in our experiment.

The main independent variable is the group-specific preferences *compatibility* indicator. Model 2 further distinguishes between groups who showed significant ($\hat{\epsilon} \geq 3$) and weak ($\hat{\epsilon} < 3$) social preferences in their other network games. Next to this, the models include six outcome-specific constants, and twelve outcome-specific measures of network *size* and network *clustering*. Together, these outcome- and network-specific variables account for many of the factors that may explain why certain investment profiles were chosen more often than others. For instance, they account for the fact that a profile was chosen because it is efficient, strategically stable, a quantal response equilibrium, or the basin of attraction of some other unobserved process. Moreover, network size and network clustering control for two of the coordination-facilitating factors found in previous experiments (Charness, Feri, Meléndez-Jiménez, and Sutter, 2014; Cassar, 2007; Berninghaus, Ehrhart, and Keser, 2002). The models additionally control for group

Table 3: Test of Hypothesis 1—Multinomial logit estimations

	refined ORE			non-refined ORE		
	$\chi = 0$ (1)	$0 < \chi < 3$ (2)	$3 \leq \chi$ (3)	$\chi = 0$ (4)	$0 < \chi < 3$ (5)	$3 \leq \chi$ (6)
Model 1:						
<i>Compatibility</i>	1.11 (0.22)	1.28 (0.45)	1.13 (0.42)	-0.59 (0.83)	0.76 (0.29)	base outcome
<i>Experience</i>	0.15 (0.09)	0.17 (0.06)	0.14 (0.09)	0.10 (0.20)	0.01 (0.06)	–
Likelihood ratio tests (<i>Compatibility</i> =0):						
(1) versus (rest)	2.12					
(1–2) versus (rest)		6.67***				
(1–3) versus (rest)			16.69***			
Model 2:						
<i>Compatibility</i> ($\hat{\epsilon} \geq 3$)	0.59 (0.35)	0.78 (0.53)	0.88 (0.38)	-0.45 (0.55)	-0.16 (0.74)	–
<i>Compatibility</i> ($\hat{\epsilon} < 3$)	1.45 (0.36)	1.61 (0.55)	1.32 (0.56)	-0.65 (1.27)	1.19 (0.41)	–
<i>Experience</i>	0.15 (0.09)	0.17 (0.05)	0.15 (0.09)	0.11 (0.20)	-0.01 (0.06)	–
Likelihood ratio tests:						
<i>Compatibility</i> ($\hat{\epsilon} \geq 3$)=0:						
(1) versus (rest)	0.10					
(1–2) versus (rest)		0.48				
(3) versus (rest)			6.26**			
<i>Compatibility</i> ($\hat{\epsilon} < 3$)=0:						
(1) versus (rest)	2.87*					
(1–2) versus (rest)		8.96***				
(3) versus (rest)			1.91			

NOTES: Coefficients and standard errors (in parentheses) of two multinomial logit models. 840 observations from final decision moments in all network games but the games on the circle. Both models include additional measures of network size and clustering (unreported). Likelihood ratio tests report χ^2 -statistics. *** $p(\chi^2) < 0.01$, ** $p(\chi^2) < 0.05$, * $p(\chi^2) < 0.1$.

experience with a certain network game because we expected that more experienced groups will find it easier to coordinate their investments on, for instance, a refined ORE.

Our hypothesis tests are based on post-estimation likelihood ratio tests, where we compared the full models against constrained models with the coefficient of *compatibility* set to zero. In both the full and the constrained models, we additionally equalized the coefficients of several outcomes to the coefficient of the base outcome (6). This way, we tested Hypothesis 1 in several different ways. For instance, the likelihood ratio test (1–3) *versus* (rest) tells us something about whether a group with compatible preferences played the refined ORE in the widest sense (any χ) more often than any other profile. The test (1) *versus* (rest), in turn, focuses on the refined ORE that are indistinguishable from a pure payoff-maximizing equilibrium.

The findings in Table 3 largely confirm Hypothesis 1. According to the

likelihood ratio tests belonging to Model 1, groups with compatible social preferences played the refined ORE more often than a non-refined ORE: χ^2 -statistic=16.69 in the test (1–3) versus (rest) significant at $p < 0.01$. They also played the refined ORE in the narrower sense (with $\chi < 3$) more often: χ^2 -statistic=6.67 significant at $p < 0.01$. Yet, they did not generally play the narrowest class of refined ORE more often than any other investment profile: χ^2 -statistic=2.12 in the test (1) versus (rest) is insignificant.

This last finding does not come entirely as a surprise. As we saw in Table 2, there is a number of subject groups who showed considerable fairness concerns and who are, therefore, expected to deviate much from the ORE that are indistinguishable from a payoff-maximizing equilibrium. We therefore estimated Model 2 and performed additional likelihood ratio tests. Much in line with our expectations, group behavior is consistent with the social preference estimated for its members. According to the likelihood ratio tests, groups with compatible and strong social preferences tend to coordinate on the refined ORE where investments deviate significantly from a payoff-maximizing best response. Groups with compatible but small social preferences chose, in contrast, the refined ORE with smaller or no deviations.

Thus, altogether, the findings in Table 2 support Hypothesis 2 that social preferences facilitate equilibrium selection. As shown in Appendix B.3, this finding is moreover robust to several alternative ways in which we can estimate our subjects' social preferences.

5.3 Hypothesis 2: network size and nestedness

The second prediction of our theory is that coordination on a refined ORE is inhibited by the size of a network and it is facilitated by the degree of nestedness, in particular the extent to which the nested neighborhoods are centralized around a few players. We provide two pieces of evidence on this.

Placebo test: The first piece of evidence comes from the circle network. According to our theory, social preferences should not help to coordinate on either a distributed or a specialized equilibrium because no neighborhood is nested in this network. To put this to a test, we searched for groups playing the circle who match the preference requirements that “worked” in other networks. Concretely, we searched for groups who match the preference requirements for the complete network and asked whether they are the ones who played the frequently observed distributed profiles on the circle. Likewise, we checked whether groups who match the preference requirements for the line network are also the ones who played the frequent specialized investment profile.

Our findings from multinomial logit estimations similar to the ones in

Table 4: Test of Hypothesis 2—Pairwise comparisons

Share of ORE with any χ					
	Star	Core	D-box	Line	Complete
	0.99	1.00	0.90	0.85	0.008
<i>Difference:</i>					
Core	-0.008				
D-box	0.09***	0.10***			
Line	0.13***	0.14***	-0.04		
Complete	0.98***	0.99***	0.89***	0.85***	
Share of ORE with $\chi < 3$					
	Star	Core	D-box	Line	Complete
	0.33	0.43	0.16	0.37	0.008
<i>Difference:</i>					
Core	-0.10				
D-box	0.17***	0.28***			
Line	-0.03	0.07	-0.20***		
Complete	0.17***	0.43***	0.15***	0.35***	
Share of ORE with $\chi = 0$					
	Star	Core	D-box	Line	Complete
	0.16	0.18	0.08	0.09	0.008
<i>Difference:</i>					
Core	-0.02				
D-box	0.08*	0.09**			
Line	0.06	0.08*	0.00		
Complete	0.15***	0.17***	0.07***	0.08***	

NOTES: Results of two-sided t-tests with 120 observations per network. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3 can be found in Appendix B.4. In line with our expectations, they suggest no systematic relationship between preference compatibility and the selection of either a distributed or specialized profile.

Hypothesis test: The second piece of evidence stems from a comparison of the numbers of refined ORE across networks. According to Hypothesis 2, we would expect a detrimental impact of network *size*. As we have already seen in Section 5.1, the share of groups who coordinated on a refined ORE was much higher on the two-player dyad network than on the four-player complete network. This difference is also statistically significant at $p < 0.001$ in a two-sided mean difference test, lending strong support to our hypothesis.

Next, Hypothesis 2 conjectured a facilitating impact of a network structure with a small number of players who nest other players' neighborhoods and a large number of players whose neighborhoods are nested. Table 4 reproduces the shares of refined ORE from Table 1 for the networks compared in Hypothesis 2. Consistent with the hypothesis, the shares of refined ORE are highest on the star and the core periphery, intermediate on the d-box and the line, and lowest on the complete network. Moreover, the two-sided mean difference tests reported in Table 4 by and large confirm the pairwise network orderings predicted in (10)–(11). The single exception is the comparison between the d-box and the line, where contrary to our theory, we find a significantly larger share of refined ORE (with $\chi < 3$) on the line net-

Table 5: Test of Hypothesis 2—OLS estimations

Model:	Share of refined other-regarding equilibria								
	(1)	any χ (2)	(3)	(4)	$\chi < 3$ (5)	(6)	(7)	$\chi = 0$ (8)	(9)
<i>One-way nestedness</i>			0.97*** (0.022)			0.42*** (0.047)			0.16*** (0.030)
<i>Nestedness</i>		1.18*** (0.011)	0.20*** (0.016)		0.45*** (0.046)	0.01 (0.024)		0.19*** (0.041)	0.03* (0.015)
<i>Clustering</i>	-0.35*** (0.019)	-0.91*** (0.023)		-0.20*** (0.046)	-0.41*** (0.032)		-0.06*** (0.012)	-0.15*** (0.027)	
<i>Constant</i>	0.74*** (0.012)	-0.01 (0.006)	-0.01 (0.006)	0.29*** (0.037)	0.01 (0.011)	0.01 (0.011)	0.11*** (0.028)	-0.01 (0.007)	-0.01 (0.007)
Observations	48	48	48	48	48	48	48	48	48
R-squared	0.09	0.80	0.85	0.12	0.51	0.53	0.03	0.27	0.29

NOTES: Results of nine OLS estimations with 48 observations each: one observation per network except the dyad (6 networks) times number of sessions (8 sessions). Standard errors in parentheses: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

work. Nevertheless, apart from this exception, Table 4 strongly supports the hypothesized conducive impact of a centralized network structure for coordination.

This gains additional support from Table 5. There, we regress the refined ORE shares for all our 4-player networks on three network statistics: first, a measure of network nestedness,

$$\text{nestedness} = \frac{1}{|N|} \sum_{i \in N} \frac{|N_i^n|}{|N_i|},$$

where N_i^n denotes the subset of i 's neighbors whose neighborhoods are nested by i or who nest i . Second, we include a measure of unidirectional nestedness,

$$\text{one-way nestedness} = \frac{1}{|N|} \sum_{i \in N} \frac{|N_i^{n*}|}{|N_i|},$$

where N_i^{n*} is the set of neighbors who are either nested by i or who nest i , but *not* both at the same time. One-way nestedness is thus high if one or a few players nest all other players' neighborhoods in a network who are just connected to these central players. We finally add clustering in some models as a control variable, since clustering explained coordination in prior experiments (e.g., Cassar, 2007; Charness, Feri, Meléndez-Jiménez, and Sutter, 2014).¹⁴ The major finding of Table 5 is that nestedness has the predicted positive effect on coordination. This effect disappears if one-way nestedness is included, as expected as well. Finally, clustering has a negative impact on coordination in our context.

¹⁴Due to the high joint correlation with nestedness and one-way nestedness, clustering is excluded from the regression models whenever the other two are included.

6 Conclusion

We set out to study how social preferences shape behavior in a network of interdependent social interactions. Towards this end, we developed a local public goods game with socially concerned players and tested the model's predictions in an experiment.

One of the main findings of this paper is that social preferences do not produce more equitable or more efficient investments in all networks alike. Rather, they help individuals to coordinate their choices in a way that reinforces the inequality that is already inherent in a network structure. In other words, if a network is fully connected, social preferences support equitable outcomes, and if a network structure is asymmetric, they support unequal payoff distributions. But, we do not find any indication that individuals would increase their investments significantly beyond what is payoff optimal.

As suggested by our theory and supported in our experiment, the central mechanism behind this finding is that socially concerned players want to maintain a certain payoff ordering in the local neighborhood of their network. In the small-scale networks of our experiment, many player groups indeed managed to coordinate on this payoff ordering. An open question from this study is, however, whether the same logic also extends to larger networks. We, therefore, leave it for future studies to advance and test our predictions for larger networks.

References

- ALLOUCH, N. (2015): "On the private provision of public goods on networks," *Journal of Economic Theory*, 157, 527–552. 2
- BELHAJ, M., S. BERVOETS, AND F. DEROÏAN (2016): "Efficient networks in games with local complementarities," *Theoretical Economics*, 11(1), 357–380. 1
- BELLEMARE, C., S. KRÖGER, AND A. VAN SOEST (2008): "Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities," *Econometrica*, 76(4), 815–839. 2, 3.2, 20
- BERNINGHAUS, S. K., K.-M. EHRHART, AND C. KESER (2002): "Conventions and local interaction structures: experimental evidence," *Games and Economic Behavior*, 39(2), 177–205. 5.2
- BERNINGHAUS, S. K., K.-M. EHRHART, AND M. OTT (2006): "A network experiment in continuous time: The influence of link costs," *Experimental Economics*, 9(3), 237–251. 4
- BINMORE, K. (2005): *Natural justice*. Oxford University Press. 2
- BOLTON, G. E., AND A. OCKENFELS (2000): "ERC: A theory of equity, reciprocity, and competition," *American Economic Review*, pp. 166–193. 3.4.5
- BONCINELLI, L., AND P. PIN (2012): "Stochastic stability in best shot network games," *Games and Economic Behavior*, 75(2), 538–554. 2

- BOURLÈS, R., Y. BRAMOULLÉ, AND E. PEREZ-RICHET (2017): “Altruism in networks,” *Econometrica*, 85(2), 675–689. 2, 6
- BRAMOULLÉ, Y., AND R. KRANTON (2007): “Public goods in networks,” *Journal of Economic Theory*, 135(1), 478–494. 1, 2, 3.1, 3.1, 7, 3.2, 3.4.5, 4, B.1, B.2
- BRAMOULLÉ, Y., R. KRANTON, AND M. D’AMOURS (2014): “Strategic interaction and networks,” *American Economic Review*, 104(3), 898–930. 2
- BRUHIN, A., E. FEHR, AND D. SCHUNK (2019): “The many faces of human sociality: Uncovering the distribution and stability of social preferences,” *Journal of the European Economic Association*, 17(4), 1025–1069. 3.2, 5.2, B.2
- CALLANDER, S., AND C. R. PLOTT (2005): “Principles of network development and evolution: An experimental study,” *Journal of Public Economics*, 89(8), 1469–1495. 4
- CASSAR, A. (2007): “Coordination and cooperation in local, random and small world networks: Experimental evidence,” *Games and Economic Behavior*, 58(2), 209–230. 5.2, 5.3
- CHARNESS, G., F. FERI, M. A. MELÉNDEZ-JIMÉNEZ, AND M. SUTTER (2014): “Experimental games on networks: Underpinnings of behavior and equilibrium selection,” *Econometrica*, 82(5), 1615–1670. 2, 5.2, 5.3, B.1, B.1
- CHARNESS, G., AND M. RABIN (2002): “Understanding social preferences with simple tests,” *The Quarterly Journal of Economics*, 117(3), 817–869. 1, 3.2, 7, B.2, B.2
- DUFWENBERG, M., AND A. PATEL (2017): “Reciprocity networks and the participation problem,” *Games and Economic Behavior*, 101, 260–272. 2, 6
- ECKEL, C. C., AND H. HARWELL (2015): “FOUR CLASSIC PUBLIC GOODS EXPERIMENTS: A REPLICATION STUDY,” *Replication in Experimental Economics*, p. 13. 1
- EYSTER, E., K. MADARÁSZ, AND P. MICHAILLAT (2021): “Pricing under fairness concerns,” *Journal of the European Economic Association*, 19(3), 1853–1898. 2
- FALK, A., A. BECKER, T. DOHMEN, B. ENKE, D. HUFFMAN, AND U. SUNDE (2018): “Global evidence on economic preferences,” *The Quarterly Journal of Economics*, 133(4), 1645–1692. 2, 3.2, 5.2, B.2
- FALK, A., AND M. KOSFELD (2012): “It’s all about connections: Evidence on network formation,” *Review of Network Economics*, 11(3). C.1
- FEHR, E., AND K. M. SCHMIDT (1999): “A theory of fairness, competition, and cooperation,” *The Quarterly Journal of Economics*, 114(3), 817–868. 3.2, 3.4.5, 9
- FEHR, E., AND I. SCHURTENBERGER (2018): “Normative foundations of human cooperation,” *Nature Human Behaviour*, 2(7), 458–468. 2
- FISCHBACHER, U. (2007): “z-Tree: Zurich toolbox for ready-made economic experiments,” *Experimental Economics*, 10(2), 171–178. 4.1, C.1
- GALEOTTI, A., S. GOYAL, M. O. JACKSON, F. VEGA-REDONDO, AND L. YARIV (2010): “Network games,” *The Review of Economic Studies*, 77(1), 218–244. 2
- GHIGLINO, C., AND S. GOYAL (2010): “Keeping up with the neighbors: social interaction in a market economy,” *Journal of the European Economic Association*, 8(1), 90–119. 2, 6
- GOYAL, S., S. ROSENKRANZ, U. WEITZEL, AND V. BUSKENS (2017): “Information acquisition and exchange in social networks,” *The Economic Journal*, 127(606), 2302–2331. 4

- GREINER, B. (2004): “An online recruitment system for economic experiments,” . C.1
- (2015): “Subject pool recruitment procedures: organizing experiments with ORSEE,” *Journal of the Economic Science Association*, 1(1), 114–125. 4.1
- HARSANYI, J. C., AND R. SELTEN (1988): “A general theory of equilibrium selection in games,” *MIT Press Books*, 1. 18
- IMMORLICA, N., R. KRANTON, M. MANEA, AND G. STODDARD (2017): “Social status in networks,” *American Economic Journal: Microeconomics*, 9(1), 1–30. 2, 6
- KERSCHBAMER, R., AND D. MÜLLER (2020): “Social preferences and political attitudes: An online experiment on a large heterogeneous sample,” *Journal of Public Economics*, 182, 104076. 2, 20
- KÖNIG, M. D., C. J. TESSONE, AND Y. ZENOU (2014): “Nestedness in networks: A theoretical model and some applications,” *Theoretical Economics*, 9(3), 695–752. 1
- MARIANI, M. S., Z.-M. REN, J. BASCOMPTE, AND C. J. TESSONE (2019): “Nestedness in complex networks: observation, emergence, and implications,” *Physics Reports*, 813, 1–90. 1, 3
- MCKELVEY, R. D., AND T. R. PALFREY (1995): “Quantal response equilibria for normal form games,” *Games and Economic Behavior*, 10(1), 6–38. B.1
- REUBEN, E., AND A. RIEDL (2013): “Enforcement of contribution norms in public good games with heterogeneous populations,” *Games and Economic Behavior*, 77(1), 122–137. 2
- RICHEFORT, L. (2018): “Warm-glow giving in networks with multiple public goods,” *International Journal of Game Theory*, 47(4), 1211–1238. 2, 4
- ROSENKRANZ, S., AND U. WEITZEL (2012): “Network structure and strategic investments: An experimental analysis,” *Games and Economic Behavior*, 75(2), 898–920. 2, 4, 13, B.1, B.1
- SELTEN, R. (1991): “Properties of a measure of predictive success,” *Mathematical social sciences*, 21(2), 153–167. 5.1, B.1
- SOBEL, J. (2005): “Interdependent Preferences and Reciprocity,” *Journal of Economic Literature*, 43(2), 392–436. 3.2
- ZHANG, Y. (2018): “Social Preferences in Network Games: Theory and Laboratory Evidence,” *Available at SSRN 3059710*. 2

A Theoretical Appendix

A.1 Proof of Lemma 1

Lemma 1 repeated. Suppose that players’ utilities are defined by the social preference function (2) with parameters $\tau_i = (\rho_i, \sigma_i, R_i)$ and the payoff function (1). A player’s social preference strength, ϵ_i , is defined by

$$\begin{aligned} \text{altruist or social-welfare types} &: \epsilon_i^p \\ \text{inequity-averse types} &: \max\{\epsilon_i^p; \epsilon_i^n\} \\ \text{competitive or spiteful types} &: \epsilon_i^n, \end{aligned}$$

where

$$\begin{aligned}\epsilon_i^p &= (b')^{-1}\left(\frac{c}{1+\rho_i}\right) - (b')^{-1}(c) \\ \epsilon_i^n &= (b')^{-1}(c) - (b')^{-1}\left(\frac{c}{1+\sigma_i}\right).\end{aligned}$$

Proof. We prove the statement for the more general case of an incomplete information game.

For a given type $\tau_i = (\rho_i, \sigma_i, R_i)$, we aim to determine an upper bound for the difference between that player's best-response investment and a payoff-maximizing best response, over all possible e_{-i} and all possible networks g . Formally, the deviation-maximizing investment $e_i = f_i(\tau_i, e_{-i})$ satisfies

$$\epsilon_i = \max \left\{ |f_i(\tau_i, e_{-i}) - f_i(e_{-i})| : \forall e_{-i} \in \mathbb{R}_+^{|\Omega_{-i}|}, \forall g \in G \right\}.$$

To find e_i and ϵ_i , let us focus on the set of locally 'tree-like' networks where no two players in i 's reference group are connected, i.e., $g_{jk} = 0 \forall j, k \in R_i$. This gives us some freedom in the design of the critical g and e_{-i} . Moreover, we focus on symmetric profiles e_{-i} , such that (i) all players j in i 's neighborhood invest the same, (ii) all players k in i 's reference group who are not in i 's neighborhood invest the same, and (iii) all players l and m in the 'hinterlands' of j and k who are not in player i 's reference group invest the same. That is, we focus on profiles e_{-i} where for all $\omega_{-i} = (\tau_1, \dots, \tau_{i-1}, \tau_{i+1}, \dots, \tau_n) \in \Omega_{-i}$: (i) $e_{\tau_j} = e_{\tau_{j'}}$ for all $j, j' \in N_i$ and $\tau_j, \tau_{j'} \in T_j$, (ii) $e_{\tau_k} = e_{\tau_{k'}}$ for all $k, k' \in R_i \setminus N_i$ and $\tau_k, \tau_{k'} \in T_k$, (iii) $e_{\tau_l} = e_{\tau_{l'}}$ for all $l, l' \notin R_i$, $l, l' \in N_j$ and $\tau_l, \tau_{l'} \in T_l$, and (iv) $e_{\tau_m} = e_{\tau_{m'}}$ for all $m, m' \notin R_i$, $m, m' \in N_k$ and $\tau_m, \tau_{m'} \in T_m$. As will become clear, the symmetry assumption, as well as the focus on tree-like networks, is innocuous as it is anyway satisfied by the deviation-maximizing e_{-i} and g defined below. Nevertheless, because utility function (2) is not continuous at investment profiles where the payoff of player i is identical to the payoff of another player j , we need to make a case distinction.

Deviation-maximizing interior solutions: Suppose first that $\pi_i(e_i, e_{-i}) \neq \pi_j(e_i, e_{-i})$ and $\pi_i(e_i, e_{-i}) \neq \pi_k(e_{-i})$. Then, i 's best-response investment, $e_i = f_i(\tau_i, e_{-i})$, must satisfy the first-order condition:¹⁵

$$b'(e_i + |N_i|e_j) - c + \rho_i \frac{|N_i|}{|R_i|} b'(e_i + e_j + |N_j|e_l) \leq 0 \quad \text{if } \pi_i > \pi_j \quad (\text{A.1})$$

$$b'(e_i + |N_i|e_j) - c + \sigma_i \frac{|N_i|}{|R_i|} b'(e_i + e_j + |N_j|e_l) \leq 0 \quad \text{if } \pi_i < \pi_j. \quad (\text{A.2})$$

The corresponding first-order condition of a payoff maximizer is

$$b'(e_i + |N_i|e_j) - c \leq 0. \quad (\text{A.3})$$

Note first that because $\rho_i \geq \sigma_i$, the maximum positive deviation, $e_i - f_i(e_{-i}) > 0$, is obtained in the range where $\pi_i > \pi_j$, and the maximum negative deviation, $e_i - f_i(e_{-i}) < 0$, in the range where $\pi_i < \pi_j$. Moreover, the absolute value of these deviations is maximal when $N_i = R_i$. We thus get that the deviation-maximizing e_i must satisfy

$$b'(e_i + |N_i|e_j) - c + \rho_i \cdot b'(e_i + e_j + |N_j|e_l) = 0 \quad \text{if } \rho_i \geq \sigma_i \geq 0 \quad (\text{A.4})$$

$$b'(e_i + |N_i|e_j) - c + \sigma_i \cdot b'(e_i + e_j + |N_j|e_l) \leq 0 \quad \text{if } \sigma_i \leq \rho_i \leq 0, \quad (\text{A.5})$$

Moreover, e_i must satisfy one of the two conditions when $\sigma_i < 0 < \rho_i$.

¹⁵Note that the relationship between π_i and π_k is irrelevant in (A.1) and (A.2) because marginally increasing e_i does not have an impact on π_k .

When we relax the assumptions of symmetric investments and locally tree-like networks, we get to the same conditions after some additional steps. The first-order condition corresponding to (A.1) and (A.2) is given by

$$\sum_{\omega_{-i} \in \Omega_{-i}} p(\omega_{-i}) \left[b'(e_i + \sum_{j \in N_i} e_{\tau_j}) - c \right. \\ \left. + \frac{\sigma_i}{|R_i|} \sum_{j \in N_i^-} b'(e_i + e_{\tau_j} + \sum_{l \in N_j \setminus \{i\}} e_{\tau_l}) + \frac{\rho_i}{|R_i|} \sum_{j \in N_i^+} b'(e_i + e_{\tau_j} + \sum_{l \in N_j \setminus \{i\}} e_{\tau_l}) \right] \leq 0,$$

where N_i^+ (N_i^-) denote the sets of neighbors with $\pi_i \geq (<) \pi_j$. Again, we want to determine the critical e_{-i} and g that maximize $|f_i(\tau_i, e_{-i}) - f_i(e_{-i})|$. Because the size of this deviation depends on the sign and the magnitude of the expressions in line two, set e_{τ_j} and e_{τ_l} such that either $\pi_i \geq \pi_j$ or $\pi_i < \pi_j$ for all types of every player j and l . Next, replace the e_{τ_l} of a player l who is neighbor of both j and i by an equally valued $e_{\tau_{l'}}$ of a player l' with $l' \notin R_i$, $l' \in N_j$. This leaves the magnitude and the sign of the deviation-maximizing e_i unaffected. Finally, choose $e_{\tau_j} = e_j$, $e_{\tau_l} = e_l$, and $|N_j| = |N_j'|$ for all $j, j' \in N_i$ so as to maximize $b'(e_i + e_{\tau_j} + \sum_{l \in N_j \setminus \{i\}} e_{\tau_l})$. In other words, the deviation-maximizing e_i involves symmetric investments e_{-i} and a tree-like network g , such as in (A.4) and (A.5).

Continuing with the main argument, note that $f_i(e_{-i}) - f_i(\tau_i, e_{-i})$ weakly increases when we lower a single neighbor's e_j in (A.5) to zero. This is because when the inequality in (A.5) is strict and the inequality in (A.3) is an equality, any reduction in e_j increases the e_i that solves (A.3) while it leaves the e_i in (A.5) unaffected. When (A.5) is an equality, instead, a reduction in e_j does not alter $f_i(e_{-i}) - f_i(\tau_i, e_{-i})$ because e_j and e_i are perfect strategic substitutes in both (A.3) and (A.5), i.e., it holds $\partial e_i / \partial e_j = -1$.

Define $b_i'' \equiv b''(e_i + |N_i|e_j)$ and $b_j'' \equiv b''(e_i + e_j + |N_j|e_l)$. The total derivatives of (A.4) and (A.5), when (A.5) has an equality sign, result in

$$\frac{\partial e_i}{\partial e_l} = \frac{-\rho_i b_j''}{b_i'' + \rho_i b_j''} < 0 \quad \text{if } \rho_i \geq \sigma_i \geq 0 \\ \frac{\partial e_i}{\partial e_l} = \frac{-\sigma_i b_j''}{b_i'' + \sigma_i b_j''} \quad \text{if } \sigma_i \leq \rho_i \leq 0.$$

Hence, e_l and e_i are strategic substitutes in equation (A.4). Moreover, because the benefit function satisfies $b''(x) \leq b''(y)$ for $x \leq y$, we can also unambiguously sign the slope of the best-response function in (A.5), which is positive. Thus, for the deviation-maximizing e_i , set $e_l = 0$ in both equations (A.4) and (A.5).

Now, because $e_j = e_l = 0$ in (A.5), the requirement $\pi_i < \pi_j$ means that the deviation-maximizing e_i must still be positive. (A.5) must thus be satisfied with equality and the maximum negative deviation is given by

$$\epsilon_i^n = (b')^{-1}(c) - (b')^{-1}\left(\frac{c}{1 + \sigma_i}\right). \quad (\text{A.6})$$

Furthermore, even though (A.4) is only valid for $\pi_i > \pi_j$, the fact that e_j and e_i are perfect strategic substitutes in both (A.3) and (A.4) allows us to reduce e_j in (A.4) to zero in order to obtain the following simple expression for the maximum positive deviation

$$\epsilon_i^p = (b')^{-1}\left(\frac{c}{1 + \rho_i}\right) - (b')^{-1}(c). \quad (\text{A.7})$$

Deviation-maximizing corner solutions: Suppose next that e_i is such that $\pi_i(e_i, e_{-i}) = \pi_j(e_i, e_{-i})$ or $\pi_i(e_i, e_{-i}) = \pi_k(e_{-i})$. Because of the discontinuity of utility function (2) at this point, the first-order conditions above do not help us any longer. Instead, we have to compare the utility of a deviation-maximizing e_i directly with the utility of the best alternative investment.

First, consider a player with $\rho_i \geq 0$ and an investment e_i with $e_i > f_i(e_{-i})$ and $\pi_i(e_i, e_{-i}) = \pi_j(e_i, e_{-i})$. This requires that for all $e'_i \neq e_i$ it holds that

$$\begin{aligned} \pi_i(e_i, e_{-i}) + \frac{|N_i|}{|R_i|} \rho_i \pi_j(e_i, e_{-i}) + \frac{|R_i| - |N_i|}{|R_i|} (\rho_i r_{ik} + \sigma_i s_{ik}) \pi_k(e_{-i}) &\geq \quad (\text{A.8}) \\ \pi_i(e'_i, e_{-i}) + \frac{|N_i|}{|R_i|} (\rho_i r'_{ij} + \sigma_i s'_{ij}) \pi_j(e'_i, e_{-i}) + \frac{|R_i| - |N_i|}{|R_i|} (\rho_i r'_{ik} + \sigma_i s'_{ik}) \pi_k(e_{-i}). \end{aligned}$$

Note first that we only need to consider deviations e'_i with $f_i(e_{-i}) \leq e'_i < e_i$. This is because (i) we are seeking the largest e_i that satisfies condition (A.8) and because (ii) any $e'_i < f_i(e_{-i})$ yields a lower utility than $e'_i = f_i(e_{-i})$.

This has two implications: First, it is $r'_{ij} = 1$ because $\pi_i(e'_i, e_{-i}) > \pi_j(e'_i, e_{-i})$ for $f_i(e_{-i}) \leq e'_i < e_i$. Second, the difference between the left-hand side (LHS) and the right-hand side (RHS) of (A.8) is maximal when $r'_{ik} = r_{ik}$ and $s'_{ik} = s_{ik}$. This is because suppose $r_{ik} = 1$ (i.e., $\pi_i(e_i, e_{-i}) \geq \pi_k(e_{-i})$). Then, $r'_{ik} = 1$. And if $s_{ik} = 1$, then $s'_{ik} \in \{0, 1\}$. So, in order to maximize the difference between LHS and RHS, and hence to maximize $|e_i - f_i(e_{-i})|$, set $r'_{ik} = r_{ik}$ and $s'_{ik} = s_{ik}$. Inequality (A.8) can then be simplified to

$$\pi_i(e_i, e_{-i}) + \frac{|N_i|}{|R_i|} \rho_i \pi_j(e_i, e_{-i}) \geq \pi_i(e'_i, e_{-i}) + \frac{|N_i|}{|R_i|} \rho_i \pi_j(e'_i, e_{-i}). \quad (\text{A.9})$$

Obviously, the above arguments also apply when we start from heterogeneous investments for player i 's neighbors and non-neighbors and a more general network structure. First, we only need to consider downward deviations e'_i from e_i for the same reasons as given above. Second, for the same reasons as above, the payoffs of i 's non-neighbors have no impact on the deviation-maximizing e_i . Third, because $\pi_j(e_i, e_{-i}) > \pi_j(e'_i, e_{-i})$ for all neighbors $j \in N_i$, let $\pi_j(e_i, e_{-i})$ enter the utility of i with a positive weight $\rho_i > 0$ to obtain a deviation-maximizing e_i . Finally, choose e_j and e_l to maximize $\pi_j(e_i, e_{-i}) - \pi_j(e'_i, e_{-i})$ for all $j \in N_i$. This leads to condition (A.9).

Continuing with the main argument, note that because $\pi_j(e_i, e_{-i}) > \pi_j(e'_i, e_{-i})$, the difference between the LHS and the RHS in (A.9) is maximal for $N_i = R_i$. Note also that $\pi_j(e_i, e_{-i}) - \pi_j(e'_i, e_{-i})$ can be written as

$$\pi_j(e_i, e_{-i}) - \pi_j(e'_i, e_{-i}) = \int_{e'_i + e_j + |N_j|e_l}^{e_i + e_j + |N_j|e_l} b'(x) dx.$$

The integral is maximal when $e_l = 0$ because $b(\cdot)$ is concave. Thus, the difference between the LHS and the RHS of (A.9) is maximal when $e_l = 0$. This means that the deviation-maximizing investment is defined by the largest e_i so that i 's utility under e_i , $[1 + \rho_i] [b(e_i + |N_i|e_j) - ce_i]$, satisfies for all e'_i with $f_i(e_{-i}) \leq e'_i < e_i$:

$$\begin{aligned} [1 + \rho_i] [b(e_i + |N_i|e_j) - ce_i] &\quad (\text{A.10}) \\ &\geq b(e'_i + |N_i|e_j) - ce'_i + \rho_i [b(e'_i + e_j) - ce_j]. \end{aligned}$$

For comparison, i 's utility under the deviation-maximizing e_i in (A.7) is given by

$$[1 + \rho_i] b(e_i) - ce_i,$$

with $e_i = (b')^{-1}(c/(1 + \rho_i))$. That term is identical to the RHS of (A.10) for $e_j = 0$ and smaller than the RHS for any $e_j > 0$.¹⁶ This means that the e_i defined in (A.10) is associated with a weakly higher utility than the e_i defined in (A.7). Therefore, the e_i in (A.10) also yields a smaller deviation from a payoff-maximizing best response

¹⁶To see why the term is smaller when $e_j > 0$, note that the e'_i that maximizes the RHS of (A.10) satisfies $e'_i : b'_i + \rho_i b'_j - c = 0$. The derivative of the RHS of (A.10) with respect to e_j is, on the other hand, given by $(b'_i + \rho_i b'_j - \rho_i c)/(1 + \rho_i)$. Combined with the first-order condition above, that derivative is thus positive since $\rho_i < 1$.

than the e_i in (A.7).

Consider next the same player with $\rho_i \geq 0$, but consider a best-response investment with $e_i > f_i(e_{-i})$ and $\pi_i(e_i, e_{-i}) = \pi_k(e_{-i})$. Then, the same arguments from above can be applied so that the resulting deviation-maximizing e_i is not larger than the e_i defined in (A.7).

Consider again the same player with $\rho_i \geq 0$, but suppose now an investment with $e_i < f_i(e_{-i})$ and $\pi_i(e_i, e_{-i}) = \pi_k(e_{-i})$. Then, i can increase her utility by deviating to $e'_i = e_i + \varepsilon$, where $\varepsilon > 0$. Hence, e_i cannot be a best-response investment.

Consider again the same player with $\rho_i \geq 0$ and an investment with $e_i < f_i(e_{-i})$. But suppose now that $\pi_i(e_i, e_{-i}) = \pi_j(e_i, e_{-i})$. Note first that for e_i to be a best response, we additionally need that $\pi_i(e'_i, e_{-i}) < \pi_j(e'_i, e_{-i})$ for all $e'_i > e_i$. The reason is the same as above that player i could otherwise increase her payoff and that of her neighbors by deviating upward. Hence, suppose that $\pi_i(e'_i, e_{-i}) < \pi_j(e'_i, e_{-i})$ for all $e'_i > e_i$ in addition. Then, e_i is a best response if for all e'_i , with $e'_i > e_i$, it holds that

$$\begin{aligned} \pi_i(e_i, e_{-i}) + \frac{|N_i|}{|R_i|} \rho_i \pi_j(e_i, e_{-i}) + \frac{|R_i| - |N_i|}{|R_i|} (\rho_i r_{ik} + \sigma_i s_{ik}) \pi_k(e_{-i}) \\ \geq \pi_i(e'_i, e_{-i}) + \frac{|N_i|}{|R_i|} \sigma_i \pi_j(e'_i, e_{-i}) + \frac{|R_i| - |N_i|}{|R_i|} (\rho_i r'_{ik} + \sigma_i s'_{ik}) \pi_k(e_{-i}). \end{aligned} \quad (\text{A.11})$$

Note now that we can apply the argument from below inequality (A.8) that the difference between the LHS and the RHS is maximal when $r'_{ik} = r_{ik}$ and $s'_{ik} = s_{ik}$. Inequality (A.11) thus simplifies to

$$\pi_i(e_i, e_{-i}) + \frac{|N_i|}{|R_i|} \rho_i \pi_j(e_i, e_{-i}) \geq \pi_i(e'_i, e_{-i}) + \frac{|N_i|}{|R_i|} \sigma_i \pi_j(e'_i, e_{-i}). \quad (\text{A.12})$$

The difference between the LHS and the RHS of (A.12) obviously depends on the size of $\pi_j(e'_i, e_{-i})$ and the sign of σ_i . Note however that $\pi_j(e'_i, e_{-i})$ can be written as

$$\pi_j(e'_i, e_{-i}) = \pi_i(e_i, e_{-i}) + \int_{e_i + e_j + |N_j|e_l}^{e'_i + e_j + |N_j|e_l} b'(x) dx. \quad (\text{A.13})$$

Thus, since $b(\cdot)$ is concave, $\pi_j(e'_i, e_{-i})$ is maximal when $e_l = 0$ and minimal when e_l is large. Consider first a player with $\rho_i \geq 0$ and $\sigma_i < 0$. Then, the difference between the LHS and the RHS in (A.12) is maximal for $e_l = 0$ and $N_i = R_i$. The deviation-maximizing e_i is thus defined by the smallest e_i so that i 's utility satisfies for all $e'_i > e_i$:

$$\begin{aligned} [1 + \rho_i] [b(e_i + |N_i|e_j) - ce_i] \\ \geq b(e'_i + |N_i|e_j) - ce'_i + \sigma_i [b(e'_i + e_j) - ce_j]. \end{aligned} \quad (\text{A.14})$$

For comparison, i 's utility under the deviation-maximizing e_i in (A.6) is given by

$$[1 + \sigma_i] b(e_i) - ce_i, \quad (\text{A.15})$$

with $e_i = (b')^{-1}(c/(1 + \sigma_i))$. That term is identical to the RHS of (A.14) for $e_j = 0$ and smaller for any $e_j > 0$.¹⁷ Therefore, we can use the same argument from below inequality (A.10) that the e_i defined in (A.14) leads to a smaller absolute deviation from a payoff-maximizing best response than the e_i in (A.6).

Consider next a player with $\rho_i \geq \sigma_i > 0$ and start from payoff expression (A.13), again. Because the RHS of (A.12) is increasing in $\pi_j(e'_i, e_{-i})$, choose a large e_l so that $\pi_j(e'_i, e_{-i}) \approx \pi_i(e_i, e_{-i})$. Moreover, set $N_i = R_i$. Inequality (A.12) can then be

¹⁷To see why the term is smaller when $e_j > 0$, note that the investment e'_i that maximizes the RHS of (A.15) needs to satisfy $b'_i + \sigma_i b'_j - c = 0$. The derivative with respect to e_j is, on the other hand, given by $(b'_i + \sigma_i b'_j - \sigma_i c)/(1 + \sigma_i)$. This derivative is thus positive.

written as

$$[1 + \rho_i] \pi_i(e_i, e_{-i}) \geq b(e'_i + |N_i|e_j) - ce'_i + \sigma_i \pi_i(e_i, e_{-i}). \quad (\text{A.16})$$

For comparison, i 's utility under the deviation-maximizing e_i in (A.6) (with $\sigma_i < 0$) can be written as in (A.15). That term is smaller than the RHS of (A.16), and thus we can again conclude that the e_i defined in (A.16) leads to a smaller maximal deviation than the e_i defined in (A.6).

We next turn to the deviation-maximizing investment of a player with $\rho_i < 0$. Consider first an investment with $e_i > f_i(e_{-i})$ and either $\pi_i(e_i, e_{-i}) = \pi_k(e_{-i})$ or $\pi_i(e_i, e_{-i}) = \pi_j(e_i, e_{-i})$. Then, i could decrease her investment to $e'_i = e_i - \varepsilon$ to raise her own payoff and lower the payoffs of her neighbors. Thus, e_i cannot be a best response.

Consider finally an investment with $e_i < f_i(e_{-i})$ and either $\pi_i(e_i, e_{-i}) = \pi_j(e_i, e_{-i})$ or $\pi_i(e_i, e_{-i}) = \pi_k(e_{-i})$. Then, we can apply the arguments as of inequality (A.11) to conclude that this leads to no larger deviation from a payoff maximizing best response than the e_i defined in (A.6). ■

A.2 Proof of Proposition 1

Proposition 1 repeated. Consider two players i and j in a network such that i nests the neighborhood of j . Moreover, suppose that either information assumption (A1) or (A2) applies. In a refined ORE, there exists at least one type combination ω such that the type of player i (j) earns weakly more (strictly less) than at least one other player k (l) in the player's reference group. That is,

$$\pi_i(\omega) \geq \min_{k \in R_i} \{\pi_k(\omega)\} \quad \text{OR} \quad \pi_j(\omega) < \max_{l \in R_j} \{\pi_l(\omega)\}$$

for at least one $\omega \in \Omega^* = T_1 \times \dots T_i^* \times \dots T_j^* \times \dots T_n$.

Proof for the complete information case. Under complete information, each player has a commonly known type and Ω^* thus reduces to a singleton. Suppose now that, contrary to the statement, player i (j) earns strictly less (weakly more) than all the players in her reference group. This implies in particular that $\pi_i(e) < \pi_j(e)$.

Combined with preference compatibility and network nestedness, we arrive at the following chain of necessary conditions for this to be a refined ORE:

$$e_i + \sum_{k \in N_i} e_k \leq f_i(e_{-i}) + \sum_{k \in N_i} e_k = e^* \leq f_j(e_{-j}) + \sum_{l \in N_j} e_l \leq e_j + \sum_{l \in N_j} e_l.$$

The first and the final inequality follow from the fact player i feels exploited or player j feels guilty (or both), so that i (j) under-(over-)invests compared to a payoff maximizer who would decide in the player's stead. The identity $f_i(e_{-i}) + \sum_{k \in N_i} e_k = e^*$ is the first-order condition of a payoff maximizer in i 's position who, due to the facts that $\pi_i(e) < \pi_j(e)$ and that i has access to more investments than j , must make a positive investment. The remaining inequality is the first-order condition of a payoff maximizer in j 's position.

Because players i and j have compatible preferences, either the first or the last inequality in the chain must be strict. This however leads to a contradiction to the requirement that i nests the neighborhood of j , which is

$$e_i + \sum_{k \in N_i} e_k \geq e_j + \sum_{l \in N_j} e_l.$$

Payoffs must thus be ordered as stated in the proposition. ■

Proof for the incomplete information case. Suppose that, contrary to the statement, all types of player i (j) earn strictly less (weakly more) under all $\omega \in \Omega^*$

than all players in their respective reference group. That is, suppose that

$$\pi_i(\omega) < \min_{k \in R_i} \{\pi_k(\omega)\} \quad \text{and} \quad \pi_j(\omega) \geq \max_{l \in R_j} \{\pi_l(\omega)\} \quad \forall \omega \in \Omega^* .$$

This immediately implies that $\pi_i(\omega) < \pi_j(\omega) \forall \omega \in \Omega^*$. Moreover, the fact that player j 's neighborhood is nested in player i 's implies that for all $\omega \in \Omega^*$:

$$e_{\tau_i} + \sum_{k \in N_i} e_{\tau_k} \geq e_{\tau_j} + \sum_{l \in N_j} e_{\tau_l} .$$

In combination, we thus require that $e_{\tau_i} > e_{\tau_j}$ for all $\omega \in \Omega^*$ because i has access to more investments than j . Yet, because i 's and j 's preferences are compatible, it must also be that

$$e_{\tau_i} \leq f_i(e_{-i}) \quad \text{and} \quad f_j(e_{-j}) \leq e_{\tau_j} \quad \forall \omega \in \Omega^* , \quad (\text{A.17})$$

with at least one inequality being strict because either player i feels exploited or player j feels guilty (or both). So, in order to simultaneously have $\pi_i(\omega) < \pi_j(\omega)$ and $e_{\tau_i} > e_{\tau_j}$, we require that $f_i(e_{-i}) > f_j(e_{-j})$.

The best-response investments of payoff maximizers in i 's and j 's position are, however, determined by the first-order conditions

$$\begin{aligned} \frac{\partial \mathbb{E} \pi_i}{\partial e_i} &= \sum_{\omega_{-i} \in \Omega_{-i}^*} p(\omega_{-i}) b' \left(f_i(e_{-i}) + \sum_{k \in N_i \setminus j} e_{\tau_k} + e_{\tau_j} \right) - c = 0 \quad (\text{A.18}) \\ \frac{\partial \mathbb{E} \pi_j}{\partial e_j} &= \sum_{\omega_{-j} \in \Omega_{-j}^*} p(\omega_{-j}) b' \left(f_j(e_{-j}) + \sum_{l \in N_j \setminus i} e_{\tau_l} + e_{\tau_i} \right) - c \leq 0 . \end{aligned}$$

Now, because i nests the neighborhood of j and because the investments of i and j are ordered as in (A.17), it follows for all $\omega \in \Omega^*$ that

$$\sum_{k \in N_i \setminus j} e_{\tau_k} + e_{\tau_j} \geq f_j(e_{-j}) + \sum_{l \in N_j \setminus i} e_{\tau_l} .$$

Combined with the concavity of $b(\cdot)$, this however means that $e_{\tau_i} > f_i(e_{-i})$ must hold for *at least one* $\omega \in \Omega^*$ in order for the first-order conditions (A.18) to be satisfied simultaneously. Hence, we arrive at a contradiction to the ordering of investments in (A.17). Payoffs must thus be ordered as stated in the proposition. ■

A.3 Experimental predictions

The following predictions characterize the refined ORE for the seven networks in Figure 1 under the perfect information assumption (A2).

A.3.1 Dyad and complete network

Suppose that, contrary to the statement, there are two players i and j with $e_i < e_j$. This means that $\pi_i(e) \geq \pi_k(e)$ for all $k \neq i$, with at least one inequality being strict. Moreover, $\pi_j(e) \leq \pi_l(e)$ for all $l \neq j$, with at least one inequality being strict. In particular, it is $\pi_i(e) > \pi_j(e)$.

Suppose now that the preferences of all players are compatible in addition. For the dyad, this means that for all $\omega \in \Omega^* = T_1^* \times T_2^*$ it holds that ω contains (i) no player of the type {altruist, spite}, (ii) no two payoff maximizers, and (iii) no two distinct types from the set {payoff maximizer, social welfare, competitive}. For the complete network, this means that all players are competitive.

Combined with the payoff inequalities shown above, preference compatibility leads to following chain of necessary conditions for $e_i < e_j$ to occur in a refined

ORE:

$$e_j + \sum_{l \in N_j} e_l \leq f_j(e_{-j}) + \sum_{l \in N_j} e_l = e^* \leq f_i(e_{-i}) + \sum_{k \in N_i} e_k \leq e_i + \sum_{k \in N_i} e_k.$$

The first and the last inequality follow from the fact player j feels exploited or player i feels guilty (or both). The identity $f_j(e_{-j}) + \sum_{l \in N_j} e_l = e^*$ is the first-order condition of a payoff maximizer in i 's position who, due to the fact that $\pi_j(e) < \pi_i(e)$, must make a positive investment. The remaining inequality is the first-order condition of a payoff maximizer in i 's position.

Because i and j have compatible preferences, either the first or the last inequality in the chain must be strict. In case of the competitive players on the complete network, it is the first inequality. This however leads to a contradiction to the fact that on the dyad and the complete network it must hold that

$$e_i + \sum_{k \in N_i} e_k = e_j + \sum_{l \in N_j} e_l.$$

Thus, it must be $e_i = e_j = e$ for all $i, j \in N$. The upper and lower bounds $e \in [(e^* \pm \epsilon)/n]$ follow, finally, from Lemma 1. ■

A.3.2 Star, core periphery, and d-box

We first apply Lemma 1 to show that an ORE on these networks entails either a center- or a periphery-sponsored public good. We then demonstrate that a center-sponsored public good *cannot* be a refined ORE.

ORE with limited preference strength: Suppose first that $e_i = 0$ for all $i \in C$ (periphery sponsorship). A payoff maximizer in the periphery position $j \in P$ of a star, core periphery, or d-box would respond with $f_j(e_{-j}) = e^*$. By Lemma 1, a social type therefore responds with $e_j = f_j(\tau_j, e_{-j})$ such that

$$e^* - \epsilon \leq e_j \leq e^* + \epsilon.$$

Moreover, social types in the duo position of the core periphery invest $e_k = f_k(\tau_k, e_{-k})$ with

$$e^* - \epsilon \leq \sum_{k \in D} e_k \leq e^* + \epsilon.$$

Next, suppose that $e_i > 0$ for at least one $i \in C$ (center sponsorship). A payoff maximizer in this position would invest

$$f_i(e_{-i}) = e^* - \sum_{l \in N \setminus i} e_l.$$

Payoff maximizers in the non-center positions $j \in P$ and $k \in D$ would, on the other hand, invest

$$\begin{aligned} f_j(e_{-j}) &= e^* - \sum_{i \in C} e_i \quad \text{and} \\ \sum_{k \in D} f_k(e_{-k}) &= e^* - e_i. \end{aligned}$$

Combined with Lemma 1, this means that the investments of social types are constrained by

$$e^* - \sum_{l \in N \setminus i} e_l - \epsilon \leq e_i \leq e^* - \sum_{l \in N \setminus i} e_l + \epsilon, \quad (\text{A.19})$$

$$e^* - \sum_{i \in C} e_i - \epsilon \leq e_j \leq e^* - \sum_{i \in C} e_i + \epsilon, \quad (\text{A.20})$$

$$e^* - e_i - \epsilon \leq \sum_{k \in D} e_k \leq e^* - e_i + \epsilon. \quad (\text{A.21})$$

It follows from (A.19) that $\sum_{i \in N} e_i \leq e^* + \epsilon$ and from (A.20) and (A.21) that

$$\begin{aligned} \sum_{i \in C} e_i + e_j &\geq e^* - \epsilon \quad \text{and} \\ e_i + \sum_{k \in D} e_k &\geq e^* - \epsilon. \end{aligned}$$

In combination, we find that the periphery players in the star and the d-box, except for periphery player p_1 , jointly contribute at most

$$\sum_{j \in P \setminus p_1} e_j = \sum_{j \in P} e_j + \sum_{i \in C} e_i - \left[\sum_{i \in C} e_i + e_{p_1} \right] \leq e^* + \epsilon - [e^* - \epsilon] = 2\epsilon.$$

Drawing the same comparison for any other periphery player p_2 , we again find that $\sum_{j \in P \setminus p_2} e_j \leq 2\epsilon$. Hence, the total contribution received by the center(s) is at most

$$\sum_{j \in P} e_j \leq \sum_{j \in P \setminus p_1} e_j + \sum_{j \in P \setminus p_2} e_j \leq 4\epsilon.$$

Similar, in the core periphery, we find for the periphery player and the duo players, respectively, that

$$\begin{aligned} e_j &= \sum_{l \in N \setminus i} e_l + e_i - [e_i + e_j] \leq e^* + \epsilon - [e^* - \epsilon] = 2\epsilon, \\ \sum_{k \in D} e_k &= \sum_{l \in N \setminus i} e_l + e_i - [e_i + \sum_{k \in D} e_k] \leq e^* + \epsilon - [e^* - \epsilon] = 2\epsilon. \end{aligned}$$

The total contribution received by the center player in the core periphery is thus, again, at most $\sum_{k \in D} e_k + e_j < 4\epsilon$. For the peripheral player with the lowest contribution in the star, the core periphery, and the d-box, (A.20) requires that $\min_j \{e_j\} + \sum_{i \in C} e_i \geq e^* - \epsilon$. Thus, the centers players' investments are necessarily larger than

$$\sum_{i \in C} e_i \geq e^* - \epsilon - \min_j \{e_j\} \geq e^* - \epsilon - \frac{4\epsilon}{|N| - |C|},$$

whereby the lower bound is determined by a situation where all peripheral (and duo) players equally share 4ϵ . Moreover, (A.19) implies that the center players' investments are necessarily smaller than

$$\sum_{i \in C} e_i \leq e^* + \epsilon.$$

Together, this defines the investment boundaries in a center-sponsored public good summarized in Table 6.

Refined ORE: In a refined ORE, it must be

$$\pi_i(e) \geq \min_{j \in N \setminus C} \{\pi_j(e)\} \quad \text{for all } i \in C. \quad (\text{A.22})$$

Suppose, to the contrary, that $\pi_i(e) < \min_{j \in N \setminus C} \{\pi_j(e)\}$ for at least one $i \in C$. We first show that in this case it must hold that $e_i \leq f_i(e_{-i})$ for all $i \in C$ and $e_j \geq f_j(e_{-j})$ for at least one $j \in P$.

To see this, suppose in addition that some or all players compare their payoffs with everyone else in the network, i.e., $R_i = N \setminus \{i\}$. Then, on the star and the core-periphery, $\tau_i \in T_i^*$ (defined in Definition 1) combined with $\pi_i(e) < \min_{j \in N \setminus C} \{\pi_j(e)\}$ implies that $e_i \leq f_i(e_{-i})$. Moreover, there is a $j \in N \setminus C$ who earns more than i and weakly more than any other $k \in N \setminus C$. When $\tau_j \in T_j^*$ for all $j \in N \setminus C$, then it certainly holds for this j that $e_j \geq f_j(e_{-j})$.

On the d-box, $\pi_i(e) < \min_{j \in N \setminus C} \{\pi_j(e)\}$ for one $i \in C$ does not necessarily imply that $j \in P$ earns less than both center players. Nevertheless, when $\tau_i \in T_i^{***}$

and $\tau_j \in T_j^{***}$ for all $i \in C$ and all $j \in P$, where

$$\begin{aligned} T_i^{***} &= \{\text{comp., spite}\} \text{ and } T_j^{***} = \{\text{payoff, social welfare, altruist}\} \\ &\quad \text{OR} \\ T_i^{***} &= \{\text{payoff, comp., spite}\} \text{ and } T_j^{***} = \{\text{social welfare, altruist}\}, \end{aligned} \tag{A.23}$$

then $e_i \leq f_i(e_{-i})$ for both $i \in C$, $e_j \geq f_j(e_{-j})$ for at least one $j \in P$, and

$$P(\omega \in \Omega^{star}) = P(\omega \in \Omega^{core}) \geq P(\omega \in \Omega^{dbox}),$$

which is what we were aiming to show.

Suppose next that *all* four players compare with their direct neighbors only, i.e., $R_i = N_i$. The requirement on the number of players with compatible social preferences can be significantly relaxed in this case. On the star and core-periphery network, $e_i \leq f_i(e_{-i})$ for the central player again requires that $\tau_i \in T_i^*$ for this player. Similar on the d-box, $e_i \leq f_i(e_{-i})$ for both $i \in C$ again requires that $\tau_i \in T_i^{***}$ for both of them. However, in the non-center positions, it suffices to have one $j \in P$ with $\tau_j \in T_j^*$ (respectively $\tau_j \in T_j^{***}$ on the d-box) for $e_j \geq f_j(e_{-j})$ because player j will only compare herself with the center player(s). The duo players in the core periphery will behave accordingly, that is, at least one of them invests $e_k \geq f_j(e_{-k})$, if $\tau_k \in T_j^*$ for both of them. Thus, again, we arrive at $e_i \leq f_i(e_{-i})$ for both $i \in C$ and $e_j \geq f_j(e_{-j})$ for at least one $j \in P$. Different to above, we have however

$$P(\omega \in \Omega^{star}) \geq \left\{ P(\omega \in \Omega^{core}) ; P(\omega \in \Omega^{dbox}) \right\}.$$

Now, because i has access to more investments than j and because $\pi_i(e) < \pi_j(e)$, we also have that $e_i > e_j$. Hence, we arrive at the following chain of conditions that need to hold if $\pi_i(e) < \min_{j \in N \setminus C} \{\pi_j(e)\}$ ought to occur in a refined ORE:

$$e_i + \sum_{k \in N \setminus i} e_k \leq f_i(e_{-i}) + \sum_{k \in N \setminus i} e_k = e^* \leq f_j(e_{-j}) + \sum_{l \in N_j} e_l \leq e_j + \sum_{l \in N_j} e_l,$$

whereby either the first or the last inequality must be strict (or both). This however leads to a contradiction to the requirement that i nests the neighborhood of j , that is, a contradiction to

$$e_i + \sum_{k \in N_i} e_k \geq e_j + \sum_{l \in N_j} e_l.$$

Thus, payoffs must be ordered as in (A.22).

Note now that in a *center-sponsored* public good, the investments of the center player(s) $i \in C$ converge to

$$\lim_{\epsilon \rightarrow 0} e_i = e^*.$$

In contrast, the investments of all other players $j \in N \setminus C$ converge to

$$\lim_{\epsilon \rightarrow 0} e_j = 0.$$

Thus, there exists an $\bar{\epsilon}^h$ such that for all $\epsilon < \bar{\epsilon}^h$ the payoff-ranking condition (A.22) cannot be satisfied.

The critical value $\bar{\epsilon}^h$ depends on the network structure. On the star and the core-periphery network, the center's payoff in a *center-sponsored* public good is necessarily lower than

$$\pi_i(e) \leq \max \pi_i(e) \equiv b(e^*) - c[e^* - 4\epsilon].$$

Because the center invests more than $e^* - 7\epsilon/3$ and each non-center player less

than 2ϵ , the non-centers' payoffs are, on the other hand, larger than

$$\pi_j(e) \geq \min \pi_j(e) \equiv b(e^* - \frac{7}{3}\epsilon).$$

Hence, $\bar{\epsilon}^{star} = \bar{\epsilon}^{core}$ is defined by the largest ϵ to satisfy $\max \pi_i(e) \leq \min \pi_j(e)$ or, equivalently,

$$c \geq \frac{b(e^*) - b(e^* - \frac{7}{3}\bar{\epsilon}^{star})}{e^* - 4\bar{\epsilon}^{star}}.$$

On the d-box, the lowest of the centers' payoffs in a *center-sponsorship* equilibrium is smaller than

$$\pi_i(e) \leq \max\{\min_{i \in C} \pi_i(e)\} \equiv b(e^*) - c \frac{e^* - 4\epsilon}{2}.$$

Because the centers invest more than $e^* - 3\epsilon$ and the peripheral players less than 2ϵ , the latter's payoffs are larger than

$$\pi_j(e) \geq \min \pi_j(e) \equiv \min b(e^* - 3\epsilon).$$

This means that $\bar{\epsilon}^{dbox}$ is defined by the largest ϵ to satisfy $\max\{\min \pi_i(e)\} \leq \min \pi_j(e)$ or, equivalently,

$$c \geq \frac{2b(e^*) - 2b(e^* - 3\bar{\epsilon}^{dbox})}{e^* - 4\bar{\epsilon}^{dbox}}.$$

A comparison of the critical values gives $\bar{\epsilon}^{star} = \bar{\epsilon}^{core} > \bar{\epsilon}^{dbox}$. ■

A.3.3 Line

ORE with limited preference strength: It follows from the definition of ϵ in (3) that $e_i + \sum_{j \in N_i} e_j \geq e^* - \epsilon$ must hold for all $i \in N$ in the line network. Suppose now that all players have small social preferences, specifically $\epsilon < e^*/3$. Then, each ORE falls into one of the following two classes:

$$\begin{aligned} \text{(end-sponsored):} & \quad ([e^* - 3\epsilon, e^* + \epsilon], [0, 2\epsilon], [0, 2\epsilon], [e^* - 3\epsilon, e^* + \epsilon]), \\ \text{(distributed):} & \quad ([e^* \pm \epsilon], 0, e_{m2} + e_{e2} \in [e^* \pm \epsilon]). \end{aligned}$$

To show this, fix the sequence of players in the order $e1 - m1 - m2 - e2$ and exclude out-of-equilibrium profiles:

- a) Obviously, *no* investment profile can be an ORE where three or more players invest nothing.
- b) Expanding on condition (3), there are three possible ORE constellations where two players invest nothing:

$$\begin{aligned} (i) : & \quad ([e^* \pm \epsilon], 0, 0, [e^* \pm \epsilon]), \\ (ii) : & \quad ([e^* \pm \epsilon], 0, [e^* \pm \epsilon], 0), \\ (iii) : & \quad (0, [e^* \pm \epsilon], [e^* \pm \epsilon], 0). \end{aligned}$$

Profiles (i) and (ii) are contained in the classes of ORE described above. Concerning (iii), the sum of player $m1$'s and $m2$'s investments must by (3) be weakly less than $e^* + \epsilon$. Hence, this is *not* an ORE when $2(e^* - \epsilon) > e^* + \epsilon$ and thus when $\epsilon < e^*/3$.

- c) There are two potential ORE configurations where one player invests noth-

ing:

$$\begin{aligned} (iv) : & \quad ([e^* \pm \epsilon], 0, e_{\tau m_2} + e_{\tau e_2} \in [e^* \pm \epsilon]), \\ (v) : & \quad (0, [e^* \pm \epsilon], e_{\tau m_2}, e_{\tau e_2}). \end{aligned}$$

The first one is contained in the classes of ORE described above. The second one is *not* an equilibrium when for player m_2 it holds that

$$\min \left\{ \sum_{i \in N} e_i \right\} = e^* - \epsilon + e^* - \epsilon > \max \{ e_{m_2} \} = e^* + \epsilon$$

and hence when $\epsilon < e^*/3$.

- d) When *all* players make a positive contribution, it follows from the best-response conditions of the end players that

$$e_{e_i} + e_{m_i} \in [e^* \pm \epsilon].$$

At the same time, the best-response conditions of the middle players require that

$$e_{e_i} + e_{m_i} + e_{m_j} \in [e^* \pm \epsilon].$$

Together, this gives $0 \leq e_{m_i} \leq 2\epsilon$ and $e^* - 3\epsilon \leq e_{e_i} \leq e^* + \epsilon$. Hence, we arrive at a profile that is contained in the classes of ORE described above.

Refined ORE: When every line middle player mi 's social preference is given by $\tau_{mi} \in T_i^{***}$ and every line end player ei 's preference is given by $\tau_{ei} \in T_j^{***}$, where T_i^{***} and T_j^{***} are defined in (A.23), then

$$\pi_{mi}(e) \geq \pi_{ei}(e) \quad \text{for } i \in \{1, 2\}. \quad (\text{A.24})$$

To see this, suppose to the contrary that $\pi_{mi}(e) < \pi_{ei}(e)$ for $i = 1$ or $i = 2$ or for both. Suppose moreover that either all players compare with their direct neighbors only, with everyone else in the network, or any combination of the two. In either case, when the preferences of all players satisfy $\tau_{mi} \in T_i^{***}$ and $\tau_{ei} \in T_j^{***}$, then it holds that $e_{mi} \leq f_{mi}(e_{-i})$ and $e_{ei} \geq f_{ei}(e_{-j})$ for $i \in \{1, 2\}$. Moreover

$$\left\{ P(\omega \in \Omega^{core}) ; P(\omega \in \Omega^{dbox}) \right\} \geq P(\omega \in \Omega^{line}).$$

Now, because mi has access to more investments than ei , $\pi_{mi}(e) < \pi_{ei}(e)$ however also implies that $e_{mi} > e_{ei}$ must hold. Hence, we arrive at the following chain of necessary conditions for this to be a refined ORE:

$$e_{mi} + \sum_{k \in N \setminus mi} e_k \leq f_{mi}(e_{-mi}) + \sum_{k \in N \setminus mi} e_k = e^* \leq f_{ei}(e_{-ei}) + e_{mi} \leq e_{ei} + e_{mi},$$

where either the first or the last inequality are strict (or both). This in turn leads to a contradiction to the requirement that ei nests the neighborhood of mi . Thus, payoffs must be ordered as in (A.24). ■

A.3.4 Circle

ORE with limited preference strength: Suppose that all players' social preferences are "small", i.e., $\epsilon < e^*/5$. We show that the classes of ORE resemble a *specialized* or a *fully distributed* investment profile.

Fix the sequence of players in the order $i - j - k - l$. First, suppose that $e_m > 0$ for all $m \in N$ (a fully distributed profile). Based on the best-response condition (3),

$e_m > 0$ must lie inside the interval $\underline{e} \leq e_m \leq \bar{e}$, where

$$\begin{aligned}\underline{e} + 2\bar{e} &= e^* - \epsilon, \\ \bar{e} + 2\underline{e} &= e^* + \epsilon.\end{aligned}$$

Solving these two identities and simplifying, we arrive at

$$\frac{e^*}{3} - \epsilon \leq e_m \leq \frac{e^*}{3} + \epsilon \quad \text{for all } m \in N.$$

Next, suppose that $e_i = 0$ for player i . It follows that player i 's neighbors, j and l , must make some positive investment. In particular, it must be $e_j > 0$ and $e_l > 0$ because suppose, to the contrary, that $e_j = 0$ (or $e_l = 0$, or both are equal to zero). Then, $e_k > 0$ since otherwise $e_i + e_j + e_k = 0$. In fact, we would require that simultaneously it holds $e_k \geq e^* - \epsilon$ and $e_l \geq e^* - \epsilon$. This, however, leads to a contradiction because it implies for a payoff maximizer in player k 's position that $f_k(e_{-k}) \leq e^* - (e^* - \epsilon)$. Thus, for a social player k we get $f_k(\tau_k, e_{-k}) = e_k \leq 2\epsilon$. However, this is at odds with the requirement $e_k \geq e^* - \epsilon$ when $\epsilon < e^*/3$. Thus, if $e_i = 0$ for player i , then it must be $e_j > 0$ and $e_l > 0$.

But this also implies that $e_k = 0$ because suppose, to the contrary, that $e_k > 0$. Because

$$e^* - \epsilon \leq e_j + e_k + e_l \leq e^* + \epsilon$$

and

$$e^* - \epsilon \leq e_k + e_l \leq e^* + \epsilon,$$

it follows that $e_l \leq 2\epsilon$ and $e_j \leq 2\epsilon$. This implies, however, that the total contribution received by player i is no larger than 4ϵ . Hence, for $\epsilon < e^*/5$ it is $e_j + e_l \leq 4\epsilon < e^* - \epsilon$. A contradiction to $e_i = 0$. Thus, it must be $e_k = 0$ (a specialized equilibrium). In particular, together with the equilibrium investments of j and l , we get that $(0, [e^* \pm \epsilon], 0, [e^* \pm \epsilon])$. ■

B Experimental Appendix

B.1 Alternative refinement concepts

To put our findings into perspective, this appendix compares our refined ORE predictions with those of several alternative equilibrium refinement concepts. Table 7 summarizes the predictions of the most relevant concepts:¹⁸

- *Asymptotic stability* (Bramoullé and Kranton, 2007) based on the idea that stable equilibria occur more frequently in our continuous-time experiment because a best-response dynamic leads back to them after a single mistake.
- *Efficiency* (Charness, Feri, Meléndez-Jiménez, and Sutter, 2014) based on the idea that subjects used the time we gave them to coordinate on a group welfare-maximizing equilibrium.
- *Quantal response* (logit) equilibria (McKelvey and Palfrey, 1995; Rosenkranz and Weitzel, 2012) based on the idea that subjects play a best response to the fluctuating, probabilistic choices of their neighbors.

As can be seen in Table 7, the alternative concepts do not explain our experimental findings better than our social preference theory. On the contrary,

- Efficiency performs worse than our social preference theory in most networks. The predictive power of efficiency is particularly low in the star and the core-periphery network, where it is efficient when the center player provides the public good all by himself but where such a center-sponsored profile is never observed. This is not entirely surprising. As suggested by the findings in Charness, Feri, Meléndez-Jiménez, and Sutter (2014), efficiency concerns are particularly powerful in games where equilibrium outcomes can be Pareto ranked, which is however not the case in our games with strategic substitutes.
- Asymptotic stability predicts better than efficiency, in particular in the star, the core periphery, and the d-box. But it fails to predict the empirically very relevant equal-split equilibria on the dyad.¹⁹
- Only quantal response theory comes close to our social preference theory. As shown in Rosenkranz and Weitzel (2012), the theory selects a unique payoff-maximizing equilibrium when players make no more than marginal decision errors. The resulting refined payoff-maximizing equilibria are identical to our refined ORE in all the networks mentioned above. Yet, quantal response theory tends to generate a too fine-grained selection on the equilibrium set. This leads to the situation that on the circle, quantal response theory predicts an egalitarian split of $e^* = 12$ as the unique equilibrium profile, even though a specialized equilibrium is equally relevant in the data.

For additional evidence, Table 8 calculates the hit ratios (Selten, 1991) of our social preference theory and the alternative refinement concepts. The hit ratio benchmarks the predictive power of a concept against the chance prediction. It is the ratio between the *observed* frequency of the predicted equilibrium type and the *predicted* frequency of this type assuming that all investment profiles are randomly drawn

¹⁸Risk dominance (Harsanyi and Selten, 1988) is omitted as a selection criterion because we deem it less relevant in our context. The subjects in our experiment were continuously informed about the investments of the other players so that strategic uncertainty is not a major concern for them.

¹⁹The reason is that all equilibrium profiles are stable on the dyad. To see this, start from a payoff-maximizing equilibrium profile with $e_i + e_j = 12$. Suppose player i would mistakenly reduce her investment e_i by x . A best response by player j would then lead to the different equilibrium $(e_i - x) + (e_j + x) = 12$.

from a set of potential profiles. The hit ratio thus takes into account that some equilibrium types might be more frequently observed because there is a larger number of investment profiles consistent with them so that they are more likely to emerge by pure chance. A hit ratio greater (smaller) 1.00 indicates a prediction that is more (less) successful than a chance prediction. In particular, the hit ratios in Table 8 benchmark the observed frequencies of an equilibrium type against its predicted likelihood given a random draw from the set of (extended) Nash equilibria. Accordingly, the sets of potential profiles in Table 8 are the sets of payoff-maximizing equilibria in columns 2–4, the sets of ORE profiles with $\chi < 3$ in column 5, and the ORE profiles with any χ in column 6.

Comparing the different columns in Table 8 shows that our social preference theory predicts at least as well as any of the alternative concepts in all network structures. In contrast to efficiency and asymptotic stability, the power of our theory is that it selects the “natural” equilibria in the dyad and all the asymmetric networks (star, core periphery, d-box, line), that is, an egalitarian equilibrium in the former and a periphery-sponsored public good in the latter. The value-added over quantal response theory is, in turn, that it does not rule out the co-existence of multiple, empirically relevant equilibria on the circle network.

B.2 Alternative social preference estimates

Table 9 categorizes our preferred parameter estimates from the main text for each subject’s (σ_i, ρ_i) -pair into seven distinct social preference types and three distinct strength classes. Notably, the preference *strength* estimates are much in line with our descriptive findings in Figure 6. There, we saw that the final investments in each network position did not deviate much from a pure payoff-maximizing best response. 84.8% of the estimates in Table 9 imply a moderate ($\hat{\epsilon}_i < 3$) and 74.6% a marginal ($\hat{\epsilon}_i < 1$) preference strength. This confirms the validity of our estimations. Strikingly also, the preference *type* estimates indicate the same sizable preference heterogeneity that was already found in earlier experiments (e.g., Charness and Rabin, 2002; Falk, Becker, Dohmen, Enke, Huffman, and Sunde, 2018; Bruhin, Fehr, and Schunk, 2019). Yet, there are some marked differences with regard to the most prevalent preference types. According to our estimations, a large majority of subjects is of an inequity-averse, competitive, or spiteful type. Meanwhile, only 2.1% and 4.8% show some concern for social welfare or altruism respectively, which are the most frequent types in the studies mentioned above.²⁰

We therefore also estimated our subjects’ social preferences in several alternative ways. Table 10 summarizes these estimates, which stem from three alternative utility functions and a different subset of experimental data. Column 1 replicates our preferred estimates based on utility function (2) and assuming that subjects only compare with their direct neighbors in a network ($R_i = N_i$). Columns 2–4 are based on a variant of this function. In particular, in Columns 2 and 3, we assume that subjects compare with everybody else, which is a plausible alternative assumption to make because subjects can see everyone’s payoffs on their screens.²¹ Columns 2 and 4, in turn, use a different subset of data as we have based our estimations there on the investment decisions in the early rounds of the network games, $t \in [20, 30)$. Our motivation is to alleviate the potential concern that the late-game decisions are “spoiled” by the early decisions by other players. In Columns 5 and 6, we, finally, base our estimations on the original distributive preference function of Charness and Rabin (2002). According to this function, a player’s social preference component consists of the *difference* between players’ own payoffs and the payoffs of their peers, rather than the absolute *level* of other players’ payoff

²⁰The estimates in Bellemare, Kröger, and Van Soest (2008) and Kerschbamer and Müller (2020) are, in contrast, closer to our findings.

²¹Against this assumption speaks the fact that subjects only influence their neighbors’ payoffs directly.

as in function (2).²² Again, just as in Columns 1–4, we estimated two variants of this utility function that differ with respect to how many players are included in a subject’s reference group.

Overall, the preference estimates based on the late-game decisions $[30, t^{max}]$ are preferable over those from the early decisions $[20, 30]$. The latter seem to be heavily influenced by the initial conditions, that is, the fact that subjects click up from initial zero investments, as indicated by the unreasonably high shares of spiteful subjects in Columns 2 and 4. Similarly, the assumption of a neighborhood comparison in Column 1 seems to be more reasonable to make because the alternative assumption in Column 3 yields an unreasonably high share of asocial types. For the same reason, the estimates in Column 1 are preferable over those in Columns 5 and 6, which are based on the original Charness and Rabin (2002) function and which show unreasonably high shares of asocial types. An issue with this utility function in the context of the Bramoullé and Kranton (2007) game is that—depending on the curvature of the benefit function $b(\cdot)$ —the function can rationalize a downward deviation from a payoff-maximizing best response as either an attempt to increase or to reduce payoff inequality. In particular, given the curvature of $b(\cdot)$ in our experiment, the function interprets the frequently observed downward deviations from the privately optimal level of $e^* = 12$ units in the peripheral positions of the star, core periphery, and d-box as attempts to *increase* payoff inequality when the focal player is *behind*. This is because an investment reduction by the peripheral player has a more detrimental impact on her own payoff than on the center’s payoff. As a result, the Charness and Rabin (2002) utility function classifies many decisions as motivated by a love of behindness, that is, it estimates positive values for σ_i . Accordingly, we find many asocial types, which are instead classified as spiteful or competitive types by our preferred model (2).

B.3 Hypothesis 1: additional evidence

Our results should not be affected by the way we estimate the social preference types of our subjects. Intuitively, if the measurement affects all social preference estimates alike, then it should not change our conclusion regarding the compatibility of preferences in a subject group.

In support of this, Table 11 replicates our results on Hypothesis 1 based on the three most meaningful alternatives to our preferred social preference estimates. The results of all three models lend support to the key mechanism behind our theory: Groups with compatible preferences coordinate more likely on a refined ORE and less likely on a non-refined ORE.

B.4 Hypothesis 2: Additional evidence

Here, we present the results of our Placebo test for the circle network. According to our theory, social preferences should fail to select equilibria on this network, even if all players share the same preference so that a distribute equilibrium might be expected, or every second player is competitive while every other player is an altruist so that a specialized profile is expected.

Table 12 presents our findings from four multinomial logit models. The dependent variable in all four models is a multinomial variable that categorizes the

²²In particular, we estimated the following function:

$$U_i(e_i, e_{-i}) = \pi_i + \frac{1}{|R_i|} \sum_{j \in R_s} (\rho_i r_{ij} + \sigma_i s_{ij}) (\pi_j - \pi_i) + \theta_{(e_i, i)},$$

where $\theta_{(e_i, i)}$ denotes a random player- and investment-specific utility component and

$$\begin{aligned} r_{ij} &= 1 \quad \text{if } \pi_i > \pi_j \text{ and } r_{ij} = 0 \text{ otherwise,} \\ s_{ij} &= 1 \quad \text{if } \pi_i < \pi_j \text{ and } s_{ij} = 0 \text{ otherwise.} \end{aligned}$$

round-end investment profiles on the circle into six different outcome classes. In Model 1, investments are categorized in terms of how much they resemble a distributed equilibrium, and in Models 2 and 3 in terms of how much they resemble one of the two specialized equilibria on the circle.²³ Model 4 presents a joint test of Models 1–3. The independent variable measures whether a subject group has a preference combination that matches the compatibility requirements for the complete network (Model 1) or the line network (Models 2 and 3). Concretely, for Model 1, we searched for groups where every subject matches the compatibility requirements defined in Appendix A.3.1. For Models 2 and 3, in contrast, we searched for groups where every second subject matches the compatibility requirements for a line middle player and every other subject the requirements for a line end player (see Appendix A.3.3).

Turning to the findings, the accompanying likelihood ratio tests do not show any sign of equilibrium selection on the circle. Because of the small number of groups who match the preference requirements for either the complete network or the line, many of the tests are inconclusive. If they are conclusive, however, the χ^2 -statistic is too small to reject the null hypothesis of no impact on equilibrium selection.

²³Table 6 presents the exact criteria for a distributed and a specialized equilibrium on the circle, respectively.

C Replication instructions

C.1 Experimental design

The computerized experiment was designed using the software program z-tree 3.0 (Fischbacher, 2007) and conducted in the Experimental Laboratory for Sociology and Economics (ELSE) at Utrecht University between 09.06. and 18.06.2008.

In the experiment, subjects had to invest in the production of a local public good in each of the seven network structures shown in Figure 1. In total, eight experimental sessions of approximately one-and-a-half hours length were scheduled and completed. Using the ORSEE recruitment system (Greiner, 2004), over 1,000 potential subjects were approached by e-mail to participate in the experiment. On average, 15 students participated per session, which gives 120 subjects in eight sessions.

A session consisted of seven treatments with varying order of treatments between the sessions. Each network structure represents a different treatment. Table 8 gives an overview.

Session	Ordering	Treatment						
		1	2	3	4	5	6	7
1	1	Dyads	Line	Star	Square	Core	Dbox	Complete
2	2	Complete	Dbox	Core	Square	Star	Line	Dyads
3	3	Dyads	Star	Line	Core	Square	Dbox	Complete
4	4	Complete	Dbox	Square	Core	Line	Star	Dyads
5	3	Dyads	Star	Line	Core	Square	Dbox	Complete
6	2	Complete	Dbox	Core	Square	Star	Line	Dyads
7	1	Dyads	Line	Star	Square	Core	Dbox	Complete
8	4	Complete	Dbox	Square	Core	Line	Star	Dyads

Figure 8: Order of treatments by session

General instructions were given before the start of a session (see the instructions below). In each treatment, subjects played a local public goods game on a given network structure five times, for 60 seconds on average under the same conditions. In particular, being positioned in a specific network, subjects could invest for a limited amount of time in order to improve their experimental points that were calculated based on formula (12).

The five repetitions of a treatment are called rounds, and each treatment consisted of one trial round and four payment rounds. At the beginning of a round, subjects were randomly allocated to a group together with either one or three other participants. Subjects were indicated as circles on the screen and could identify themselves by color: Each subject saw him- or herself as a blue circle while all neighbors were represented as black circles (see below for a screenshot).

Each round had the same structure and lasted between 30 and 90 seconds. The round ends were unknown and randomly determined. Starting from a situation of zero investments, subjects indicated simultaneously on their computer terminals (by clicking on one of two buttons at the bottom of the screen) whether they wished to in- or decrease their investment. Full information about the momentary investments of all other subjects was continuously provided and updated five times per second by the computer. Also, the resulting payoffs of all participants could continuously be observed on the screen. At the end of a round, subjects were informed about the number of points they earned with the investments they saw on their screen. In other words, final earnings only depended on the situation at the end of a round.

Subjects were not identifiable between different rounds or at the end of the experiment. In this fashion, we minimized the dependence across observations (Falk and Kosfeld, 2012). Taking the seven treatments together, every subject played 35 network games in 35 different groups, of which 28 were payoff relevant. Altogether, this gives 960 networks games and 3,360 investment decisions (8 sessions times 15 subjects on average per session times 7 treatments (6 networks of 4 subjects and 1 network of 2 subjects) times 4 cycles). At the end of the experiment,

the experimental points were converted into euros at a rate of 400 points = 1 euro. In addition, subjects received a 3 euro participation fee. The average earning was thus 11.82 euros.

C.2 Experimental instructions

C.2.1 English version



- Instructions -

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

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

400 points = 1 Euro

The money you earn will be paid out in cash at the end of the experiment without other participants being able to see how much you earned. Further instructions on this will follow in due time. During the experiment you are not allowed to communicate with other participants. Turn off your mobile phone and put it in your bag. Also, you may only use the functions on the screen that are necessary for the functioning of the experiment. Thank you very much.

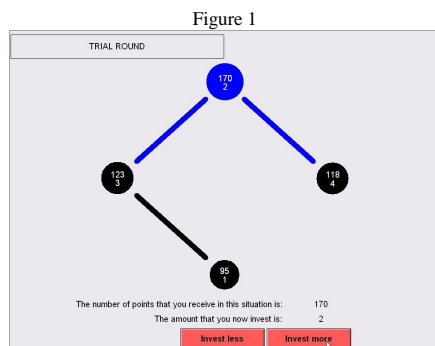
- Overview of the experiment -

The experiment consists of *seven scenarios*. Each scenario consists again of *one trial round* and *four paid rounds* (altogether 35 rounds of which 28 are relevant for your earnings).

In *all scenarios* you will be *grouped* with either one or with three other randomly selected participants. At the beginning of *each of the 35 rounds*, the groups and the positions within the groups will be randomly changed. The participants that you are grouped with in one round are very likely different participants from those you will be grouped with in the next round. It will not be revealed with whom you were grouped at any moment during or after the experiment.

The participants in your group (of two or four players, depending on the scenario) will be shown as circles on the screen (see Figure 1). You are displayed as a **blue** circle, while the other participants are displayed as **black** circles. You are always connected to one or more other participants in your group. These other participants will be called *your neighbors*. These connections differ per scenario and are displayed as lines between the circles on the screen (see also Figure 1).

Each round lasts *between 30 and 90 seconds*. The end will be at an unknown and random moment in this time interval. During this time interval you can earn **points** by producing know-how, but producing know-how also costs points. The points you receive in the end depend on your own investment in know-how and the investments of your neighbors.



By clicking on one of the two buttons at the bottom of the screen you increase or decrease your investment in know-how. At the end of the round, you receive the amount of points that is shown on the screen at that moment in time. In other words, your final earnings only depend on the situation at the end of every round. Note that this end can be at any between 30 and 90 seconds after the round is started and that this moment is unknown to everybody. Also different rounds will not last equally long.

The points you will *receive* can be seen as the *top number* in your blue circle. The points others will *receive* are indicated as the *top number* in the black circles of others. Next to this, the *size of the circles* changes with the points that you and the other participants will receive: a larger circle means that the particular participant receives more points. The *bottom number* in the circles indicates the amount *invested* in know-how by the participants in your group.

Remarks:

- It can occur that there is a time-lag between your click and the changes of the numbers on the screen. One click is enough to change your investment by one. A subsequent click will not be effective until the first click is effectuated.
- **Therefore wait until your investment in know-how is adapted before making further changes!**

- Your earnings -

Now we explain how the number of points that you earn depends on the investments. Read this carefully. Do not worry if you find it difficult to grasp immediately. We also present an example with calculations below. Next to this, there is a trial round for each scenario to gain experience with how your investment affects your points.

In all scenarios, the points you receive at the end of each round depend in the same way on two factors:

1. **Every unit that you invest in know-how yourself will cost you 5 points.**
2. **You earn points for each unit that you invest yourself and for each unit that your neighbors invest.**

If you sum up all units of investment of yourself and your neighbors, the following table gives you the points that you earn from these investments:

Your investment plus your neighbors' investments	0	1	2	3	4	5	6	7	8	9	10
Points	0	28	54	78	100	120	138	154	168	180	190

Your investment plus your neighbors' investments	11	12	13	14	15	16	17	18	19	20	21
Points	198	204	208	210	211	212	213	214	215	216	217

The higher the total investments, the lower are the points earned from an additional unit of investment. Beyond an investment of 21, you earn one extra point for every additional unit invested by you or one of your neighbors.

Note: if your and your neighbors' investments add up to 12 or more, earnings increase by less than 5 points for each additional unit of investment.

- Example -

Suppose

1. you invest 2 units;
2. one of your neighbors invests 3 units and another neighbor invests 4 units.

Then you have to pay 2 times 5 = 10 points for your own investment.

The investments that you profit from are your own plus your neighbors' investments: 2 + 3 + 4 = 9. In the table you can see that your earnings from this are 180 points.

In total, this implies that you receive $180 - 10 = 170$ points if this would be the situation at the end of the round. Figure 1 shows this example as it would appear on the screen. The investment of the fourth participant in your group does not affect your earnings. In the trial round before each of the seven scenarios, you will have time to get used to how the points you will receive change with investments.

- Scenarios -

All rounds are basically the same. The only thing that changes between scenarios is whether you are in a group of two or four participants and how participants are connected to each other. Also your own position randomly changes within scenarios and between rounds. We will notify you each time on the screen when a new scenario and trial round starts. At the top of the screen you can also see when you are in a trial round (see top left in Figure 1). Paying rounds are just indicated by "ROUND" while trial rounds are indicated by "TRIAL ROUND".

- Questionnaire -

After the 35 rounds you will be asked to fill in a questionnaire. Please take your time to fill in this questionnaire accurately. In the mean time your earnings will be counted. Please remain seated until the payment has taken place.

C.2.2 Dutch version



- Instructions -

Neemt u alstublieft de volgende instructies aandachtig door. Hierin staat alles wat u moet weten om deel te nemen aan het experiment. Indien u vragen hebt, steek uw hand op. Er zal iemand bij u komen om uw vraag te beantwoorden. Deze regels zijn hetzelfde voor alle deelnemers.

U kunt geld verdienen tijdens dit experiment door het vergaren van punten. Het aantal punten dat u verdient, hangt af van uw eigen keuzes en van de keuzes van andere deelnemers. Het totaal aantal punten dat u verdient in het experiment zal aan het einde van het experiment omgewisseld worden tegen de wisselkoers van:

400 punten = 1 Euro

Aan het einde van het experiment krijgt u het geld dat u verdient hebt tijdens het experiment contant uitbetaald. Later volgen hierover verdere instructies. Tijdens het experiment is het niet toegestaan te communiceren met andere deelnemers. Zet u mobiele telefoon uit en berg hem op in uw tas. U mag ook alleen de functies op het scherm activeren die nodig zijn voor het functioneren van het experiment. Hartelijk dank.

- Overzicht van het experiment -

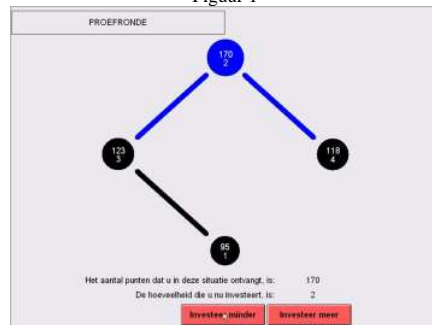
Het experiment bestaat uit *zeven scenario's*. Elk scenario bestaat weer uit *één proefronde* en *vier betaalde rondes* (samen 35 rondes waarvan er 28 relevant zijn voor uw verdiensten).

In alle scenario's wordt in een groep geplaatst met één of drie andere deelnemers. Aan het begin van *elk van de 35 rondes* worden de groepen en de posities binnen de groepen willekeurig veranderd. De deelnemers waarmee u in de ene ronde in een groep zit, zijn zeer waarschijnlijk andere deelnemers dan diegene waarmee u in de volgende ronde in een groep zit. Tijdens of na het experiment zal het niet bekend worden gemaakt met wie u in een groep gezeten hebt.

De deelnemers in uw groep (dat zijn er twee of vier afhankelijk van het scenario) worden als cirkels weergegeven op het scherm (zie Figuur 1). U wordt zelf weergegeven met een **blauwe** cirkel, terwijl de andere deelnemers worden weergegeven als **zwarte** cirkels. U bent altijd verbonden met één of meer andere deelnemers in uw groep. Deze andere deelnemers noemen we *uw burens*. Deze verbindingen verschillen per scenario and worden weergegeven met lijnen tussen de cirkels op het scherm (zie ook Figuur 1).

Elke ronde duurt *tussen de 30 en 90 seconden*. Het einde zal op een onbekend en willekeurig moment in dit tijdsinterval plaatsvinden. Tijdens dit tijdsinterval kunt u punten verdienen door kennis te produceren, maar de productie van kennis kost ook punten. De punten die u aan het einde ontvangt, hangen af van uw eigen investering in kennis en van de investeringen van uw burens.

Figuur 1



Door te klikken op de twee knoppen onder aan het scherm, kunt u uw investering in kennis verhogen of verlagen. Aan het einde van elke ronde, ontvangt u het aantal punten dat op dat moment op het scherm wordt weergegeven. Uw uitbetaling hangt dus alleen af van de situatie aan het einde van elke ronde. Merk op dat dit einde komt op een voor iedereen onbekend moment tussen de 30 en 90 seconden na het begin van de ronde. Verschillende rondes zullen ook niet even lang duren.

Het aantal punten dat u zult ontvangen zijn weergegeven als het bovenste getal in uw blauwe cirkel. De punten die anderen zullen ontvangen zijn weergegeven als het bovenste getal in hun zwarte cirkels. Daarnaast verandert *de grootte van de cirkels* met het aantal punten dat u of de andere deelnemers zullen krijgen: een groter cirkel betekent dat die deelnemer meer punten zal verdienen. Het onderste getal in de cirkels geeft het aantal punten weer dat de deelnemers in uw groep investeren in kennis.

Opmerkingen:

- Het kan gebeuren dat er een vertraging is tussen uw klik en de veranderingen van de getallen op het scherm. Eén klik is voldoende om uw investering met één punt te veranderen. Een volgende klik zal pas effect hebben als de eerste klik is verwerkt.
- **Wacht daarom met een volgende klik totdat uw eerdere verandering verwerkt is op het scherm!**

- Uw verdiensten -

Nu leggen we uit hoe uw verdiensten afhangen van de investeringen. Lees dit zorgvuldig! Wees niet bezorgd als het niet meteen helemaal duidelijk is. We zullen zodadelijk ook een rekenvoorbeeld laten zijn. Daarnaast is er bij elk scenario een proefronde om ervaring te krijgen met hoe uw investering uw aantal punten bepaalt.

In alle scenario's hangt het aantal punten dat u ontvangt aan het einde van een ronde af van twee factoren:

1. **Elke eenheid die u investeert in kennis kost uzelf 5 punten.**
2. **U verdient punten met elke eenheid die uzelf investeert en met elke eenheid die uw burens investeren.**

Als u de hoeveelheid die uzelf investeert en de investeringen van uw burens optelt, geeft de volgende tabel weer hoeveel punten u verdient met deze investeringen:

Uw investeringen plus investeringen van uw burens	0	1	2	3	4	5	6	7	8	9	10
Punten	0	28	54	78	100	120	138	154	168	180	190

Uw investeringen plus investeringen van uw burens	11	12	13	14	15	16	17	18	19	20	21
Punten	198	204	208	210	211	212	213	214	215	216	217

Hoe hoger de investeringen worden, hoe minder punten erbij komen voor nieuwe investeringen. Als het totaal van investeringen boven de 21 komt, ontvangt u nog één punt voor elke volgende eenheid die u of een van uw burens investeert.

Let op: als uw investering plus die van uw burens samen 12 of meer zijn, stijgen uw verdiensten met minder dan 5 punten per extra eenheid investering.

- Voorbeeld -

Stel

1. u investeert 2 eenheden;
2. één van uw burens investeert 3 eenheden, een andere buur 4 eenheden.

Dan moet u 2 keer 5 = 10 punten betalen voor uw eigen investering.

De investeringen waarvan u profiteert zijn uw eigen investering plus die van uw burens: 2 + 3 + 4 = 9. In de tabel kunt u zien dat dit u 180 punten oplevert.

In totaal betekent dit dat u 180 – 10 = 170 punten verdient als dit de situatie zou zijn aan het einde van de ronde. Figuur 1 laat dit voorbeeld zien zoals het op uw scherm verschijnt. De investering van de vierde deelnemer in uw groep heeft geen effect op uw aantal punten. In de proefronde aan het begin van elk scenario krijgt u de kans om te wennen aan hoe de punten die u ontvangt veranderen met de investeringen.

- Scenario's -

Alle rondes zijn in principe hetzelfde. Het enige wat verandert tussen de scenario's is de manier waarop u met andere deelnemers verbonden bent. Ook zal uw eigen positie in een groep kunnen veranderen tussen rondes. U krijgt elke keer een mededeling op het scherm als een nieuw scenario en een proefronde begint. Bovenaan het scherm kunt u ook zien of u in een proefronde zit (zie Figuur 1). Betaalde rondes worden aangegeven met alleen "RONDE", terwijl proefrondes worden aangegeven met "PROEFRONDE".

- Vragenlijst -

Aan het einde van de 35 rondes vragen we u nog om een vragenlijst in te vullen. Neem alstublieft rustig de tijd om deze vragenlijst precies in te vullen. Ondertussen tellen wij uw verdiensten. Blijft u op uw plek totdat de betaling is afgerond.

3

C.3 Selection and eligibility of participants

Subjects subscribe to a database via a website (www.elseutrecht.nl), which explains the type of experiments that they subscribe for. The Welcome-text is shown in Appendix C.4.

All subjects are recruited from this database. By subscribing a subject indicates her willingness to participate in the type of experiments described. This means that by subscribing, a subject in principle agrees to participate in the described type of task. All experiments exclusively involve computerized tasks.

At the beginning of an experiment, subjects are informed that if for any reason they might not be willing to continue, they can notify the experiment leader and stop the experiment (for details on the rules, see elseutrecht.nl/public/rules.php). No further explicit consent form is used for individual experiments.

C.4 Recruitment text

Recruitment text participants

Welcome! This is the web site of the "Experimental Laboratory for Sociology and Economics"(ELSE) at Utrecht University.

ELSE is a computer room. It is used for studying social science and economics research questions in an experimental manner. For this purpose, we are looking for people who are interested to participate in our experiments. During an experiment, up to 30 participants at a time can get a place at one of the computers in ELSE. The participants anonymously interact with other participants via the computer and answer a short questionnaire about themselves. On average, participants earn between 8 and 10 EURO per hour, but this amount can vary between studies. The exact amount of money received typically depends on the decisions made by oneself and other participants. The experiments do not involve other tasks than making decisions during anonymous interactions, and answering questions via the computer, unless this is explicitly mentioned in an invitation.

Are you interested in earning money for your decisions and answering our questions, support science and gain some insight into this research field? Then be welcome to participate in our experiments!

In order to participate, you first need to subscribe to the participants database via this web site (click on register in the menu on the left side). Every now and then you will receive a message inviting you for a specific experiment. At that moment you can decide whether you want to participate in a specific experiment (see Rules)

Before you subscribe to the participants database, you need to indicate that you agree with the rules we follow. Please read the information provided on our web site concerning:

- Participation and Rules of Proper Laboratory Behavior -
Researchers' Commitments and Privacy Policy -
Frequently Asked Questions (FAQs)

If you have further questions, please feel free to contact us using the e-mail address below.

ELSE is located at the Uithof in the Sjoerd Groenmangebouw, Padualaan 14 at the 3rd floor in room A3.03.

C.5 Approval of the Institutional Review Board

This experiment is one of a series of experiments conducted for a project entitled "Cooperation in Social and Economical relationship". The Ethics committee of Social and Behavioral Sciences of Utrecht University granted joint approval to all the experiments of this project including the current experiment. The approval was filed on October 22, 2017, under number FETC17-028.



Universiteit Utrecht

P.O. Box 80140, 3508 TC Utrecht

The Board of the Faculty of Social and Behavioural Sciences
Utrecht University
P.O. Box 80.140
3508 TC Utrecht

**Faculty of Social and
Behavioural Sciences**
Faculty Support Office
Ethics Committee

Visiting Address
Padualaan 14
3584 CH Utrecht

Our Description	FETC17-028 (Buskens)
Telephone	030 253 46 33
E-mail	FETC-fsw@uu.nl
Website	https://intranet.uu.nl/facultaire-ethische-toetsingscommissie-fetc
Date	October 22, 2017
Subject	Ethical approval

ETHICAL APPROVAL

Study: cooperation in social and economical relations

Principal investigator: prof. Vincent Buskens, Ph.D.

This research project does not belong to the regimen of the Dutch Act on Medical Research Involving Human Subjects, and therefore there is no need for approval of a Medical Ethics Committee.

The study is approved by the Ethics Committee of the Faculty of Social and Behavioural Sciences of Utrecht University. The approval is based on the documents send by the researchers as requested in the form of the Ethics committee and filed under number FETC17-028 (Buskens). Given the review reference of the Ethics Committee, there are no objections to execution of the proposed research project, as described in the protocol. It should be noticed that any changes in the research design oblige a renewed review by the Ethics Committee.

Yours sincerely,

Peter van der Heijden, Ph.D.
Chair

Jacqueline Tenkink-de Jong LL.M.
Executive secretary

Table 6: Experimental predictions

	Payoff-maximizing equilibria	Other-regarding equilibria
Dyad and complete	$\sum_{i \in N} e_i = 12$ (S,E) (Q: $e_i = e_j = \frac{12}{n}$)	$\sum_{i \in N} e_i \in [12 \pm \epsilon]$ Refined ORE: $e_i = e_j \in [\frac{12 \pm \epsilon}{n}]$
Star	(i) $e_c = 0, e_p = 12$ (ii) $e_c = 12, e_p = 0$ (S,Q: (i) selected) (E: (ii) selected)	(i) $e_c = 0, e_p \in [12 \pm \epsilon]$ (ii) $e_c \in [12 - \frac{7\epsilon}{3}, 12 + \epsilon]$, $\sum_{j \in P} e_j \leq 4\epsilon$ Refined ORE: $\pi_c \geq \min_{j \in N \setminus c} \{\pi_j\}$ Refined ORE with $\epsilon < 3$: (i) selected
Core periphery	(i) $e_c = 0, e_p = 12$, $\sum_{j \in D} e_j = 12$ (ii) $e_c = 12, e_{-c} = 0$ (S: (i) selected) (Q: (i) selected with $e_d = 6$) (E: (ii) selected)	(i) $e_c = 0, e_p \in [12 \pm \epsilon]$, $\sum_{j \in D} e_j \in [12 \pm \epsilon]$ (ii) $e_c \in [12 - \frac{7\epsilon}{3}, 12 + \epsilon]$, $\sum_{j \neq c} e_j \leq 4\epsilon$ Refined ORE: $\pi_c \geq \min_{j \in N \setminus c} \{\pi_j\}$ Refined ORE with $\epsilon < 3$: (i) selected
D-box	(i) $e_c = 0, e_p = 12$ (E) (ii) $e_p = 0, \sum_{i \in C} e_i = 12$ (E) (S,Q: (i) selected)	(i) $e_c = 0, e_p \in [12 \pm \epsilon]$ (ii) $\sum_{i \in C} e_i \in [12 - 3\epsilon, 12 + \epsilon]$, $\sum_{j \in P} e_j \leq 4\epsilon$ Refined ORE : $\pi_c \geq \min_{j \in P} \{\pi_j\}$ Refined ORE with $\epsilon < 2$: (i) selected
Line	(i) $e_{ei} = 12, e_{mi} = 0$, $e_{mj} + e_{ej} = 12$ (S) (ii) $e_{mj} = 0, e_{ej} = 12$ (Q) (iii) $e_{mj} = 12, e_{ej} = 0$ (E)	$\forall i : e_i + \sum_{j \in N_i} e_j \geq e^* - \epsilon$ ORE with $\epsilon < 3$: (i) $e_{ei} \in [12 \pm \epsilon], e_{mi} = 0$, $e_{mj} + e_{ej} \in [12 \pm \epsilon]$, (ii) $e_e \in [12 - 3\epsilon, 12 + \epsilon]$, $e_m \leq 2\epsilon$ Refined ORE: $\pi_{mi} \geq \pi_{ei}$
Circle	(i) $e_i = 0, e_{i+1} = 12$ (ii) $e_i = 4$ (S,E: (i) selected) (Q: (ii) selected)	$\forall i : e_i + e_{i-1} + e_{i+1} \geq e^* - \epsilon$ ORE with $\epsilon < 3$: (i) $e_i = 0, e_{i+1} \in [12 \pm \epsilon]$ (ii) $e_i \in [4 \pm \epsilon]$

NOTES: (Other-regarding) equilibria for the experimental games where decisions in the final round are rewarded according to payoff function (12) with $e^* = 12$. For comparison, the equilibria selected by alternative equilibrium refinement methods are indicated as well: (S) asymptotic stability, (Q) quantal response equilibria with marginal decision errors, (E) efficient equilibria.

Table 7: Frequency of refined equilibria

Network	Equilibrium type	Deviation from payoff-maximizing equilibrium		
		zero ($\chi = 0$)	moderate ($\chi < 3$)	any (any χ)
Dyad	equal	32.1% (S,E,Q,rfd)	45.8% (rfd)	49.2% (rfd)
	other	8.8% (S,E)	33.0%	50.8%
Complete	equal	0.8% (S,E,Q,rfd)	0.8% (rfd)	0.8% (rfd)
	other	20.8% (S,E)	62.5%	99.2%
Star	per-spon	15.8% (S,Q,rfd)	33.3% (rfd)	62.5% (rfd)
	cent-sp: $\pi_c \geq \pi_j$	—	—	36.6% (rfd)
	cent-sp: other	0% (E)	0.8%	0.8%
Circle	spec	7.5% (S,E)	16.6%	29.2%
	distr	3.3% (Q)	27.5%	55.0%
Core	per-spon	17.5% (S,Q,rfd)	43.3% (rfd)	68.3% (rfd)
	cent-sp: $\pi_c \geq \pi_j$	—	—	31.7% (rfd)
	cent-sp: other	0% (E)	0%	0%
D-box	per-spon	8.3% (S,E,Q,rfd)	15.0% (rfd)	25.8% (rfd)
	cent-sp: $\pi_c \geq \pi_j$	—	0.8% (rfd)	64.2% (rfd)
	cent-sp: other	0% (E)	3.3%	10.0%
Line	end-spon	0.8% (S,Q,rfd)	10.0% (rfd)	46.7% (rfd)
	distr: $\pi_m \geq \pi_e$	8.3% (S,rfd)	26.7% (rfd)	39.2% (rfd)
	distr: other	0.8% (S,E)	0.8%	14.1%

NOTES: Percentages of (refined) equilibrium profiles at the random ends of the 960 network games. 240 observations for dyad, 120 for all other networks. Refined equilibria are: (Q) quantal response, (S) stable, (E) efficient, (rfd) refined other-regarding equilibria.

Table 8: Hit ratios of alternative refinement concepts

	Efficiency	Stability	Quantal response	Refined ORE		
				$\chi = 0$	$\chi < 3$	any χ
Dyad	1.00	1.00	10.20	10.20	12.59	8.36
Complete	1.00	1.00	16.85	16.85	9.88	39.30
Star	0.00	2.00	2.00	2.00	118.39	1.16
Circle	1.04	1.04	0.48	1.00	1.00	1.00
Core periphery	0.00	1.08	1.08	1.08	17.80	1.20
D-box	1.00	14.00	14.00	14.00	40.57	1.51
Line	0.15	1.00	1.99	1.99	1.05	2.01
Average	0.60	3.02	6.66	6.73	28.75	7.79

Table 9: Preference type distribution

Preference type	Preference strength		
	any (any $\hat{\epsilon}_i$)	moderate ($\hat{\epsilon}_i < 3$)	marginal ($\hat{\epsilon}_i < 1$)
altruism ($\hat{\rho}_i \geq \hat{\sigma}_i > 0$)	4.8%	4.8%	4.8%
social welfare ($\hat{\rho}_i > \hat{\sigma}_i = 0$)	2.1%	2.1%	1.9%
inequity-aversion ($\hat{\rho}_i > 0 > \hat{\sigma}_i$)	10.2%	5.4%	3.5%
competitive ($0 = \hat{\rho}_i > \hat{\sigma}_i$)	30.6%	28.8%	25.4%
spiteful ($0 > \hat{\rho}_i \geq \hat{\sigma}_i$)	20.8%	18.8%	14.0%
payoff maximizer ($\hat{\rho}_i = \hat{\sigma}_i = 0$)	23.3%	23.3%	23.3%
asocial ($\hat{\sigma}_i > 0 > \hat{\rho}_i$)	8.1%	1.7%	1.7%
	100.0%	84.8%	74.6%

NOTES: Categorization of estimated $(\hat{\sigma}_i, \hat{\rho}_i)$ -pairs according to revealed preference type and revealed preference strength. A value of $-0.0465 < x < 0.048$ for $x \in \{\hat{\sigma}_i, \hat{\rho}_i\}$ is set to zero because a player with such a small preference parameter would take a decision indistinguishable from a payoff-maximizer. The mapping of the other $(\hat{\sigma}_i, \hat{\rho}_i)$ -pairs into types and strengths is based on utility function (2) and Lemma 1 applied to the experimental payoff function (12), yielding $\epsilon_i^p = 2.5\rho_i/(1 + \rho_i)$ and $\epsilon_i^n = 2.5\sigma_i/(1 + \sigma_i)$ as preference strength measures. All remaining estimates, including pairs with $\hat{\sigma}_i \leq -1$ or $\hat{\rho}_i \geq 1$, are categorized as *asocial* types.

Table 10: Alternative type distributions

Utility function:	Function (2)				C&R function	
	neighbors		all		neighbors	all
Reference group:						
Estimation period:	$[30, t^{max}]$	$[20, 30]$	$[30, t^{max}]$	$[20, 30]$	$[30, t^{max}]$	$[30, t^{max}]$
Type distribution:	(1)	(2)	(3)	(4)	(5)	(6)
altruist	4.8	0.6	3.1	0.8	17.5	26.9
social welfare	2.1	0.2	2.1	0.4	4.4	5.4
inequity-averse	10.2	10.6	2.9	1.7	15.2	15.6
competitive	30.6	25.4	15.8	10.4	4.2	2.3
spiteful	20.8	41.0	15.4	59.2	7.3	3.1
payoff-max.	23.3	5.4	41.9	11.9	3.3	2.1
asocial	8.1	16.7	18.8	15.6	48.1	44.6
	100	100	100	100	100	100

NOTES: Classification of estimated $(\hat{\sigma}_i, \hat{\rho}_i)$ -pairs into social preference types. Estimates are based on subject- and game-specific conditional logit estimations of four alternative utility functions and two alternative estimation periods. See Section 4.2 for procedural details.

Table 11: Multinomial logit results for alternative type distributions

	refined ORE			non-refined ORE		
	$\chi = 0$ (1)	$0 < \chi < 3$ (2)	$3 \leq \chi$ (3)	$\chi = 0$ (4)	$0 < \chi < 3$ (5)	$3 \leq \chi$ (6)
Model 1 (Type distribution 3):						
<i>Compatibility</i>	0.60 (0.57)	0.37 (0.56)	–	0.53 (0.87)	-0.05 (0.67)	-0.30 (0.68)
LR tests (<i>Compatibility</i> =0):						
(1) versus (rest)	2.25					
(1-2) versus (rest)		3.18*				
(1-3) versus (rest)			1.00			
Model 2 (Type distribution 5):						
<i>Compatibility</i>	0.87 (0.40)	0.75 (0.17)	–	-0.02 (1.10)	0.47 (0.49)	-25.14 (0.46)
LR tests (<i>Compatibility</i> =0):						
(1) versus (rest)	5.87**					
(1-2) versus (rest)		11.76***				
(1-3) versus (rest)			3.02*			
Model 3 (Type distribution 6):						
<i>Compatibility</i>	1.22 (0.78)	1.04 (0.70)	–	0.65 (0.58)	0.72 (0.89)	-16.22 (0.87)
LR tests (<i>Compatibility</i> =0):						
(1) versus (rest)	2.32					
(1-2) versus (rest)		3.85**				
(1-3) versus (rest)			1.96			

NOTES: Coefficient estimates and standard errors (clustered at the session level in parentheses) of three multinomial logit models. Models are based on 840 observations from final decision moments of all networks games, except the games on the circle. All models include an additional group-specific *experience* variable, measuring the x -th repetition of the same network game, and two measures of network *size* and *clustering*. Likelihood ratio tests report χ^2 -statistics. *** $p(\chi^2) < 0.01$, ** $p(\chi^2) < 0.05$, * $p(\chi^2) < 0.1$.

Table 12: Placebo test on the circle—Multinomial logit results

Model 1:	distributed ORE			non-distributed ORE		
	$\chi = 0$ (1)	$0 < \chi < 3$ (2)	$3 \leq \chi$ (3)	$\chi = 0$ (4)	$0 < \chi < 3$ (5)	$3 \leq \chi$ (6)
<i>Compatibility</i>	-0.26 (0.79)	-0.16 (0.28)	14.88 (1.07)	— —	-0.17 (0.32)	base outcome
Likelihood ratio test (<i>Compatibility</i> =0): (1–3) versus (rest)						
—						
Model 2:	ORE: players 2 and 4 as free riders			other ORE		
	$\chi = 0$ (1)	$0 < \chi < 3$ (2)	$3 \leq \chi$ (3)	$\chi = 0$ (4)	$0 < \chi < 3$ (5)	$3 \leq \chi$ (6)
<i>Compatibility</i>	-14.55 (0.85)	-14.52 (0.80)	-14.52 (0.94)	-14.49 (0.98)	1.06 (0.67)	base outcome
Likelihood ratio test (<i>Compatibility</i> =0): (1–3) versus (rest)						
—						
Model 3:	ORE: players 1 and 3 as free riders			other ORE		
	$\chi = 0$ (1)	$0 < \chi < 3$ (2)	$3 \leq \chi$ (3)	$\chi = 0$ (4)	$0 < \chi < 3$ (5)	$3 \leq \chi$ (6)
<i>Compatibility</i>	-12.64 (1.22)	1.36 (1.05)	1.21 (1.23)	-12.34 (1.00)	-0.17 (1.06)	base outcome
Likelihood ratio tests (<i>Compatibility</i> =0):						
(1) versus (rest) —						
(1–2) versus (rest) 0.40						
(1–3) versus (rest) 1.98						
Model 4:	distributed/specialized ORE			other ORE		
	$\chi = 0$ (1)	$0 < \chi < 3$ (2)	$3 \leq \chi$ (3)	$\chi = 0$ (4)	$0 < \chi < 3$ (5)	$3 \leq \chi$ (6)
<i>Compatibility</i>	-13.88 (0.56)	-0.48 (1.01)	0.78 (0.93)	-13.77 (0.86)	0.55 (0.64)	base outcome
Likelihood ratio tests (<i>Compatibility</i> =0):						
(1) versus (rest) —						
(1–2) versus (rest) 1.11						
(1–3) versus (rest) 0.00						

NOTES: Results of four multinomial logit estimations for the final investments on the circle network. Models 1–3: 120 observations; Model 4: pooled data of Models 1–3. All models include a group-specific experience measure (measuring the x -th repetition of the game on the circle). Likelihood ratio tests report χ^2 -statistics.

Declarations

- Funding: Funding was provided by Utrecht University
- Conflicts of interest/Competing interests: Attached below.
- Availability of data and material:

The authors confirm that the data (raw and processed) supporting the findings of this study will be available from the corresponding author on request.

- Code availability

The authors confirm that the zTree files (programs used to collect experimental data) and Stata codes (Programs used to perform the data analysis) of this study will available from the corresponding author on request.

Hagen, 23 July 2021

DISCLOSURE STATEMENT

In the past three years I did not receive financial support summing to at least \$10,000 (in the form of consultant fees, retainers, grants and the like) from an "interested" party.

In the past three years I did not hold any paid or unpaid positions as officer, director, or board member of relevant non-profit organizations or profit-making entities.

To the best of my knowledge there have been no conflict of interest including any financial, personal or other relationships with other people or organizations within three years of beginning the submitted work that could have inappropriately influenced this work.

A handwritten signature in black ink, appearing to read "B. Westbrook". The signature is written in a cursive style with a large initial "B" and a long, sweeping underline.

Bastian Westbrook

Utrecht, 23 July 2021

DISCLOSURE STATEMENT

In the past three years I received significant financial support, summing to at least \$10,000 (in the form of consultant fees, retainers, grants and the like), from the following sources:

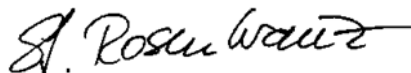
- 1) Utrecht University, Heidelberglaan 1, 3584 CS Utrecht
- 2) PBL Netherlands Environmental Assessment Agency, Bezuidenhoutseweg 30, 2594 AV The Hague
- 3) CPB Netherlands Bureau for Economic Policy Analysis, Bezuidenhoutseweg 30, 2594 AV The Hague

In the past three years I held the following paid or unpaid positions as officer, director, or board member of relevant non-profit organizations or profit-making entities:

- 1) Stichting GXP, Utrecht, Director (no compensation)

I declare that none of the above sponsors/positions influenced the academic neutrality of the content, conclusions, or implications of the submitted paper.

To the best of my knowledge there have been no conflict of interest including any financial, personal or other relationships with other people or organizations within three years of beginning the submitted work that could have inappropriately influenced this work.



Stephanie Rosenkranz

Utrecht, 23 July 2021

DISCLOSURE STATEMENT

In the last three years I, or a close relative or partner, did not receive financial support summing to at least US\$10,000 (in the form of consultant fees, retainers, grants and the like, or in kind support, such as providing access to data) from an "interested" party (any individual, group, or organization that has a financial, ideological, or political stake related to the article).

I, or a close relative or partner, did not hold any paid or unpaid positions (as officer, director, or board member of relevant non-profit organizations or profit-making entities) at a "relevant" organization, whose policy positions, goals, or financial interests relate to the article.

To the best of my knowledge there have been no conflict of interest including any financial, personal or other relationships with other people or organizations within three years of beginning the submitted work that could have inappropriately influenced this work.

A handwritten signature in black ink, appearing to read 'S. Rezaei', written in a cursive style.

Sarah Rezaei

Amsterdam, 23 July 2021

DISCLOSURE STATEMENT

In the past three years I received significant financial support, summing to at least \$10,000 (in the form of consultant fees, retainers, grants and the like), from the following sources:

- 4) Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081HV Amsterdam
- 5) Radboud University Nijmegen, Heyendaalseweg 141, 6525 AJ Nijmegen
- 6) Ministerie van Financiën, Korte Voorhout 7, 2500 EE Den Haag

In the past three years I held the following paid or unpaid positions as officer, director, or board member of relevant non-profit organizations or profit-making entities:

- 2) Society for Experimental Finance, Innsbruck, Director (no compensation)
- 3) Stichting GXP, Utrecht, Director (no compensation)

I declare that none of the above sponsors/positions influenced the academic neutrality of the content, conclusions, or implications of the submitted paper.

To the best of my knowledge there have been no conflict of interest including any financial, personal or other relationships with other people or organizations within three years of beginning the submitted work that could have inappropriately influenced this work.



Utz Weitzel