Inequality and Fairness: A Networked Experiment[1]

Ugur Yildirim[2], Dennis Feehan[3]

July 8, 2020

Abstract. Why do humans cooperate? Lab experiments have found that cooperation may emerge in part because humans have intrinsically egalitarian motives, meaning that they resist inequality even at some personal cost (Dawes et al 2007, Fehr, Bernhard, and Rockenbach 2008, Johnson et al 2009, Xiao and Bicchieri 2010). But outside the lab, economic inequality is high and on the rise (Piketty and Saez 2006, McCall and Percheski 2010, Alvaredo et al 2018), yet survey data suggest that people do not prioritize policies intended to address inequality (Dallinger 2010, McCall 2013). If people are intrinsically egalitarian, why are dramatic increases in inequality not a bigger concern? One possibility is that most people care more about unfairness than inequality *per se* (Tyler 2011, Bjornskov et al 2013, Starmans, Sheskin, and Bloom 2017). Here, we report the results of a networked, online experiment designed to unpack the relationship between fairness and inequality. In our experiment, we create fair and unfair wealth allocations by experimentally manipulating two factors: wealth *distribution* (i.e., whether starting wealth is equal vs unequal) and wealth *source* (i.e., the specific mechanism through which wealth (in)equality comes about, earned vs random). Our results show that the source of subjects' wealth has important effects on their attitudes and behavior: when subjects "earned" their endowments, they perceived their wealth regimes to be more fair, and they were less likely to cooperate. These findings suggest that it can be misleading to study inequality without accounting for subjects' understanding of how that inequality arose.

Keywords: inequality; fairness; sociology; social networks; experimental design

[1] The authors thank Daniel Schneider, David Harding, Gabriel Lenz, Sophia Rabe-Hesketh, Geoffrey Wodtke, and Xinyi Zhang for all of their helpful comments, Mark McKnight, Carl Mason, and Carl Boe for all of their technical help, without which this project could not have been undertaken, UC Berkeley XLab, Don Moore in particular, for the generous research grant that funded this study, and the anonymous study participants who agreed to be a part of our study.
[2] Ugur Yildirim (ugur.yildirim@berkeley.edu) is a Ph.D. candidate in the Department of Sociology, University of California-Berkeley.
[3] Dennis Feehan (feehan@berkeley.edu) is an assistant professor in the Department of Demography, University of California-Berkeley.

Introduction

Economic inequality is high and on the rise in many countries around the globe (e.g., Piketty and Saez 2006, McCall and Percheski 2010, Alvaredo et al 2018). Yet survey data suggest that people are not always supportive of policies intended to address inequality (Dallinger 2010, McCall 2013), and that they would generally prefer wealth distributions that are at least somewhat unequal (Norton and Ariely 2011, Norton 2014). This is puzzling, since psychologists, evolutionary biologists, economists, and others have repeatedly found evidence that people have intrinsically egalitarian motives, meaning that they would be expected to resist inequality, even at some personal cost (Dawes et al 2007, Fehr, Bernhard, and Rockenbach 2008, Johnson et al 2009, Xiao and Bicchieri 2010). If people are intrinsically egalitarian, why are dramatic increases in inequality not a bigger concern?

One possibility is that most people care more about unfairness than inequality *per se* (Tyler 2011, Bjornskov et al 2013, Starmans, Sheskin, and Bloom 2017). Although scholars have long made a conceptual distinction between fairness and equality (McCall 2013), most empirical studies have examined the simultaneous effects of these concepts (Starmans, Sheskin, and Bloom 2017). To help unpack the relationship between inequality and unfairness, we designed an experiment based on a networked iterated public goods game. Networked public goods games have been widely used as an experimental model for understanding social decision-making, and they have formed the basis for many important empirical findings about the emergence and persistence of cooperation and similar prosocial behavior. For example, recent research has used networked iterated public goods games to find that dynamic networks promote prosocial behavior (Wang, Suri, and Watts 2012, Shirado et al 2013); that, under certain conditions, static networks can promote prosocial behavior (Rand et al 2014); that reputation moderates the role of static vs dynamic networks in promoting prosocial behavior (Melamed, Harrell, and Simpson 2018); that optimal network fluidity can promote economic growth and minimize inequality (Nishi, Shirado, and Christakis 2015); and that the visibility of wealth can amplify effects of initial inequality (Nishi et al 2015).

In the classic networked iterated public goods game, subjects are linked together in a network and then asked to play several rounds of a public goods game with their network neighbors. At the start of the experiment, researchers grant subjects an initial amount of wealth, called an endowment, with which the game is to be played. In our experiment, we modify the classic game by experimentally manipulating the *endowment regime*, which is determined by two factors: endowment source (the specific mechanism through which starting endowments are chosen, earned vs random) and endowment distribution (whether starting wealth is equal or unequal). Previous studies have not experimentally manipulated the source of the subjects' endowment at the start of the game (Nishi et al 2015). But, by experimentally manipulating the source of subjects' endowments, our design enables us to estimate how subjects' reactions to inequality might be moderated by the process that produces inequality.

We use the results of our experiment to study two important questions: first, what causes an endowment regime to be perceived as fair or unfair? And, second, which endowment regimes lead to prosocial behavior? Our results show that the source of subjects' endowments has important effects: when subjects "earned" their endowments, they perceived their endowment regimes to be more fair, and they were less likely to behave prosocially. These findings suggest that it can be misleading to study inequality without accounting for subjects' understanding of how that inequality arose.

Methods

Skill-based task and endowment allocation

In our experiment, groups of 12 to 18 subjects first 'earn' points by completing a skill-based task, which is a fill-in-the-blanks word game (see Appendix 1 for screenshots). Each subject in a network is presented with a list of words that are missing some letters. Subjects are asked to identify as many incomplete words as possible in a limited amount of time. For example, if a subject is shown "t_adit_on_l", then they would earn points for responding "traditional", and no points for any other response. The more words a subject successfully identifies, the higher her score on the task. This first phase is identical for subjects in all experimental conditions.

After completing the skill-based task, each group of subjects is randomly assigned to one of four endowment regimes: (i) earned unequal (EU), (ii) random unequal (RU), (iii) earned equal (EE), or (iv) random equal (RE) (Figure 1). The earned unequal regime assigns each subject whatever score they were able to achieve in the task; the random unequal regime randomly shuffles the group of subjects' scores from the task before assigning them to subjects; the earned equal regime assigns each subject the same score, which is the average of the group of subjects' scores on the task; and the random equal regime assigns each subject in the group the same score, which is randomly picked from among previously achieved scores on the task (Figure 1).

Iterated public goods game

Subjects play several rounds of an iterated public goods game, where the initial distribution of wealth is governed by which endowment regime subjects were randomized into. Playing this public goods game gives subjects the opportunity to make a sequence of decisions that can be either prosocial or selfish. That is, subjects repeatedly make the decision to either (1) prosocially contribute to a public good; or (2) refrain from contributing to a public good (thus free-riding when others do contribute).

When a subject contributes to the public good, they pay a cost and give everyone they are connected to benefits. Thus, through the decisions made in the public goods game, each subject's stock of points can increase, decrease, or stay the same after each round. The more points a subject finishes the game with, the more they are paid at the end of the study. Subjects are not told exactly when the game will end in order to avoid endgame effects. (After each round of the public goods game, subjects may also be offered the opportunity to unlink/link themselves to other subjects so as to ensure that they are not necessarily stuck with whomever they happen to be connected at the beginning.)

After subjects finish the first public goods game, this process is then repeated a second time: the group of subjects is randomly assigned to participate in a second endowment regime whose wealth allocation rules may be different from the first. Subjects play a second public goods game according to the rules of the second endowment regime. Since respondents are randomized into one of four endowment regimes for the first game, and then again into one of four endowment regimes for the second game, there are a total of 16 different experimental conditions. Figure 1 provides a conceptual overview of the design.[4] We record subjects'

[4] See Appendix 1 for the full set of texts and figures used in the experiment.

behavioral actions, and also ask subjects survey questions intended to gauge their attitudes towards the fairness of the endowment regime that they were randomized into.

Distinctive features of our design

Our design has two distinctive features. The first distinctive feature is that we randomize the rules used to allocate subjects' endowments in the iterated public goods games along two independent dimensions: endowment source (earned vs random), and endowment distribution (equal vs unequal). Earlier experiments have not manipulated the source of the endowment (Nishi et al 2015). The advantage of manipulating endowment source in addition to endowment distribution is that it enables us to investigate how subjects' reactions to inequality might be moderated by the process that produces inequality.

The second distinctive feature of our design is that subjects play two separate public goods games, allowing us to observe how individual participants' attitudes and behaviors change from one endowment regime to another. Since the endowment regime governing each game was assigned randomly, our design enables causal inferences to be made about the effects of fairness and inequality on attitudes and behaviors. Formally, these inferences come from within-subjects models. The within-subjects approach is a significant improvement over the conventional between-subjects approach because (1) the within-subjects approach gives the study more statistical power; and (2) the within-subjects approach implicitly controls for time-invariant confounders such as age, sex, and education (which remain constant from the first game to the second).

Implementation and subject recruitment

Our experiment was implemented using the Breadboard platform (McKnight and Christakis 2016). Following several recent studies on cooperation in networked public goods games (e.g., Nishi et al 2015), subjects were recruited from Amazon's Mechanical Turk (mTurk) platform. Subjects were required to (i) be located in the US; (ii) have an overall HIT approval rate of 90% or above; and (iii) not have previously participated in the study.

A total of 1,870 mTurk Workers participated in 160 sessions, and 1,759 of them fully completed the task. There were 9 rounds in the first public goods game and 10 rounds in the second public goods game, meaning that there is complete data on a total of $1,759 \times (9+10) = 33,421$ subject-rounds.[5] The research project has IRB approval and was pre-registered on OSF. All code, materials, and de-identified data will be made public once the study is over.

Results

We discuss results from two within-subjects models that were fit to the experimental data.[6] The first model focuses on perceived fairness by examining how changes in the endowment regimes subjects were assigned to affect changes in how fair subjects report each game to be. The second model focuses on prosocial behavior by examining how changes in endowment regimes affect the probability that each subject chose to cooperate, a key behavioral outcome.[7]

[5] See Appendix 2 for details of sample size calculations, and Appendix 3 for a detailed description of all variables in the dataset, including demographic information.

[6] See Appendix 4 for details on within-subjects models, and Appendix 5 for between-subjects models.

[7] Appendix 6 shows results for two other behavioral outcomes: tie formation and tie breakage.

Perceived fairness

The left panel of Figure 2 shows average changes in perceived fairness for all 16 experimental conditions. (The figure is produced based on estimates from a model where the outcome is the change in perceived fairness from the first to the second game and the predictors are the 16 dummies corresponding to the 16 possible condition pairs, e.g., RU to EU is one possibility, without an intercept. A zero-sum constraint is applied to the model whereby the average changes in the 16 cells add up to 0.)

The Figure suggests that subjects perceived the earned unequal (EU) endowment regime to be the most fair: playing the first game in EU and the second game in any other condition leads to a decrease in perceived fairness while playing the first game in any other condition and the second game in EU leads to an increase in perceived fairness. Thus, endowment regimes with equal initial wealth distributions were not perceived to be the most fair in our study.

To estimate the causal effects of endowment source (random, earned) and endowment distribution (unequal, equal) on subjects' fairness perceptions, we fit a within-subjects model where the outcome is the change in fairness score from the first game to the second, and the predictors are the changes in endowment source, the changes in endowment distribution, and the changes in their interaction. Formally, the model fitted to estimate the effect of change in endowment regime on change in fairness perceptions can be described using Equation 1.

$$\Delta fairness_{ij} = \delta_0 + \delta_1 \Delta earned_i + \delta_2 \Delta equal_i + \delta_3 \Delta earned_i \times equal_i + e_{ij} \quad (1)$$

The variables $earned_i$, $equal_i$, and $earned_i \times equal_i$ only take the values 0 and 1, so the variables $\Delta earned_i$, $\Delta equal_i$, and $\Delta earned_i \times equal_i$ -- the changes in endowment source, endowment distribution, and their interaction -- can take the values -1, 0, or 1. For example, when $\Delta earned_i$ is 1, this means that subjects in a given network session were in a "random" condition in the first game and an "earned" condition in the second game. Similarly, since $fairness_{ij}$ can take discrete values between 1 and 7 (higher values more fair), $\Delta fairness_{ij}$ can take discrete values between -6 to 6. The interaction term is the difference of the interaction terms for each game, i.e. $\Delta earned_i \times equal_i = earned_{i2} \times equal_{i2} - earned_{i1} \times equal_{i1}$. Standard errors are clustered at the level network (i) level.[8]

This model treats changes in endowment regimes as a continuous variable. If the effect is symmetric, that is, $change\ to\ earned = -1 \times change\ to\ random$, then the continuous estimator is statistically more efficient compared to an estimator that treats changes in endowment regimes as a categorical variable. Formal tests show that it is not possible to reject the hypothesis that the effects are symmetric, so we focus on the more efficient continuous estimator (p=0.683 in the earned case, and p=0.898 in the equal case).[9]

[8] See Appendix 4 for alternative results based on a linear mixed-effects model that explicitly decomposes the error term into two parts.

[9] See Appendix 4 for results based on both types of models. The model that treats change in endowment regime as a categorical variable with the categories -1 (change to unequal/random), 0 (no change), and 1 (change to equal/earned) shows that the most salient effect is the one corresponding to a change from an earned to a random regime (e.g., EU to RU), which is negative (≈ -0.6). Regardless, the most salient effect is due to the Earned/Random rather than the Equal/Unequal axis.

Figure 3 (left panel) shows the coefficient estimates from this model. The figure shows that changing into an "earned" condition leads to a 1.386 ($p<0.001$) increase in fairness, while changing into an "equal" condition leads to a smaller 0.482 ($p=0.001$) increase in fairness; the interaction effect is large and negative (coefficient estimate=-1.706, $p<0.001$). These estimates are in line with the patterns observed in the left panel of Figure 2, in particular the salient "earned unequal" effect. The evidence strongly suggests that subjects perceived "earned unequal" to be the most fair condition. (See also Figure A5.1 for a between-subjects fairness ordering of the four endowment regimes.)

Prosocial behavior

The right panel of Figure 2 shows average changes in the probability of cooperating for all 16 experimental conditions. Patterns shown here suggest that changes into "earned" conditions lead to decreases in cooperation, while changes into "random" conditions lead to increases in cooperation. (The figure is produced based on estimates from a model where the outcome is the change in cooperation choice from the first to the second game and the predictors are the 16 dummies corresponding to the 16 possible condition pairs, e.g., RU to EU is one possibility, without an intercept, controlling for round. The model adjusts for the general drop in cooperation between games using a zero-sum constraint whereby the average changes in the 16 cells add up to 0.)

To estimate the causal effect of endowment source and endowment distribution on subjects' cooperation decisions, we again fit a within-subjects model where the outcome is the change in the decisions to cooperate from the first game to the second, and the predictors are the changes in endowment source, the changes in endowment distribution, the changes in their interaction, and dummies for the round in which the cooperation decision was made. Formally, the model fitted to estimate the effect of endowment regime on cooperation decisions can be described using Equation 2.

$$\Delta cooperation_{ijl} = \delta_0 + \delta_1 \Delta earned_i + \delta_2 \Delta equal_i + \delta_3 \Delta earned_i \times equal_i + \sum_{l=1}^{9} \delta_{4l}\, round_{il} + e_{ijl} \quad (2)$$

In this equation, the predictors $\Delta earned_i$, $\Delta equal_i$, and $\Delta earned_i \times equal_i$ have exactly the same interpretation as Equation 1 and take the values -1, 0, or 1. The outcome variable $\Delta cooperation_{ijl}$ -- the change in cooperation decision -- similarly takes the values -1, 0, or 1. Standard errors are clustered by network and subject.[10] Changes in endowment regimes are once again treated as a continuous, rather than a categorical, variable since formal tests show that it is not possible to reject the hypothesis that the effects are symmetric, suggesting that the continuous estimator is reasonable ($p=0.160$ in the earned case, and $p=0.850$ in the equal case).[11]

Figure 3 (right panel) shows the coefficient estimates from this model. The estimates show that changing into an "earned" condition leads to a -4.15 ($p=0.010$) percentage point decrease in cooperation, while the "equal" effect is close to zero and statistically insignificant

[10] See Appendix 4 for alternative results based on a linear mixed-effects model as well as models that are fit on network-level, rather than individual-level data.

[11] See Appendix 4 for results based on both types of models. The models that treat changes in endowment regimes as a categorical predictor further indicate that the main driver of this effect is a significant drop in cooperation when changing to an Earned regime.

(coefficient estimate=-0.24, p=0.881) (the interaction effect is similarly insignificant with coefficient estimate=1.47 and p=0.520). These results suggest that when players have earned their wealth, either individually or as a group, they are less willing to behave prosocially by contributing to a public good. The results are robust to several changes in model specification (see Appendix 4).

Comparing the patterns in the left and right panels of Figure 2 suggests that there is less cooperation in endowment regimes that are more fair (this inverse relationship can also be seen more directly in Figure A4.1). However, our design does not directly manipulate people's perceived fairness; instead, our experiment is designed such that subjects are first randomized into an endowment regime (EE, EU, RE, RU); next, they are asked to assign a fairness score to the condition they find themselves in; finally, they play a public goods game with other subjects in the network session. Given this flow, subjects' fairness perceptions are endogenous to the endowment regime. If we assume that changes in endowment regimes affect cooperation choices only through changes in fairness perceptions, then a direct test of the relationship between perceived fairness and cooperation can be made by using an instrumental variable model to estimate the endogenous fairness effect. The model uses changes in endowment regimes to instrument changes in fairness perceptions; that is, the model relates changes in perceived fairness that are induced as a result of random assignment to different endowment regimes to changes in subjects' propensity to cooperate. This additional model tells us that a one unit increase in fairness score leads to a -2.23 (p=0.031) percentage point decrease in cooperation.[12]

Discussion

We find that the source of subjects' endowment has important effects on the course of the game: when subjects "earned" their endowments, they perceived their endowment regimes to be more fair, and they were less likely to behave prosocially. Thus, our results suggest that it can be misleading to study inequality without accounting for subjects' understanding of how that inequality arose.

When endowments were earned, subjects in our study found unequal endowment regimes to be more fair than equal ones. This is surprising in light of previous research that has suggested that people have egalitarian motives and an aversion to inequality (e.g., Dawes et al 2007, Fehr, Bernhard, and Rockenbach 2008, Johnson et al 2009, Xiao and Bicchieri 2010). If this were true, then we would have expected the equal endowment regimes to be perceived as fair and the unequal ones to be perceived as unfair. However, our results tell a more nuanced story: while the unequal regime where initial wealth is *earned* was perceived to be the *most fair*, the unequal regime where initial wealth is *random* was perceived to be the *least fair*. Pairwise comparisons (see Appendix 5) further suggest that subjects prefer fair regimes. Our findings are thus more consistent with the hypothesis that most people care more about unfairness than inequality *per se* (Tyler 2011, Starmans, Sheskin, and Bloom 2017).

Earlier studies suggested that humans might consider randomly allocated wealth (e.g., through coin flips or lotteries) to be fair (e.g., Kimbrough, Sheremeta, and Shields 2014). Subjects in our study found randomly allocated endowments to be less fair than earned endowments, contradicting this view. Again, our results suggest a more complex picture:

[12] See Appendix 4 for more details on the instrumental variable model.

subjects perceived the regime in which endowments were randomly but unequally allocated to be the least fair of all.

Our behavioral results showed that endowment source affected prosociality: subjects chose to contribute less to the public good under earned compared to random conditions. This finding hints at a potential entitlement effect in the sense that people become less willing to take risky, prosocial actions if they feel entitled to their endowments. This interpretation is in line with earlier, mainly survey-based, evidence that shows that perceptions of individual effort, equal opportunity, or simply a general sense of fairness make people less likely to support redistribution (Alesina and Angeletos 2005, Alesina and Ferrara 2005, Bjornskov et al 2013). The results are also consistent with previous behavioral findings using alternate types of economic games: for example, researchers have found that average contributions to other group members are lower when performance in a task, rather than sheer luck, determines success (Krawczyk 2010).

Our study contributes to the literature on equality and fairness by looking at a specific source of fairness: skill (as opposed to randomness). Future work could investigate alternative sources of fairness. As suggested by Starmans, Sheskin, and Bloom (2017), "the idea that people's discontent with the current distribution of wealth has to do with fairness, rather than inequality itself, opens up a wealth of new questions about which factors (for example, hard work, skill, need, morality) are psychologically relevant for fair distributions." (p. 5) In particular, researchers might be interested in disentangling hard work/effort from skill/ability. One way to achieve this would be to design a study where endowment is determined by either a simple task purely based on effort (e.g., clicking the mouse) or a more complex task that requires some skill (e.g., trivia). Furthemore, qualitative assessments of why people act the way they do under a given endowment regime would shed further light into the reasons why people choose to make prosocial decisions. Finally, our findings are based on the US and future work could investigate whether or not this result holds in other contexts (e.g., Jakiela 2012).

Citations

Alesina, A. and G. Angeletos. 2005. "Fairness and Redistribution." *The American Economic Review* 95(4): 960-980.

Alesina, A. and E. Ferrara. 2005. "Preferences for redistribution in the land of opportunities." *Journal of Public Economics* 89: 897-931.

Alvaredo, F., L. Chancel, T. Piketty, E. Saez, G. Zucman. 2018. "The Elephant Curve of Global Inequality and Growth." *AEA Papers and Proceedings* 108: 103-108.

Bjornskov, C., A. Dreher, J. Fischer, J. Schnellenbach, and K. Gehring. 2013. "Inequality and Happiness: When Perceived Social Mobility and Economic Reality Do Not Match." *Journal of Economic Behavior and Organization* 91: 75-92.

Dallinger, U. 2010. "Public support for redistribution: what explains cross-national differences?" *Journal of European Social Policy* 20(4): 333-349.

Dawes, C., J. Fowler, T. Johnson, R. McElreath, and O. Smirnov. 2007. "Egalitarian motives in humans." *Nature* 446: 794-796.

Fehr, E., H. Bernhard, and B. Rockenbach. 2008. "Egalitarianism in young children." *Nature* 454: 1079-1083.

Jakiela, Pamela. 2012. "Social Preferences and Fairness Norms as Informal Institutions: Experimental Evidence." *American Economic Review* 101(3): 509-513.

Johnson, T., C. Dawes, J. Fowler, R. McElreath, and O. Smirnov. 2009. “The role of egalitarian motives in altruistic punishment.” *Economics Letters* 102(3): 192-194.

Kimbrough, E., R. Sheremeta, and T. Shields. 2014. “When Parity Promotes Peace: Resolving Conflict Between Asymmetric Agents.” *Journal of Economic Behavior & Organization* 99: 96-108.

Krawczyk, M. 2010. “A glimpse through the veil of ignorance: Equality of opportunity and support for redistribution.” *Journal of Public Economics* 94: 131-141.

McCall, L. 2013. *The Undeserving Rich American Beliefs about Inequality, Opportunity, and Redistribution.* New York: Cambridge University Press.

McCall, L. and C. Percheski. 2010. “Income Inequality: New Trends and Research Directions.” *Annual Review of Sociology* 36: 329-347.

McKnight, M. and N. Christakis. *Breadboard*. Computer software. *Breadboard: Software for Online Social Experiments.* Vers. 2. Yale University, 1 May 2016. Web.

Melamed, D., A. Harrell, and B. Simpson. 2018. “Cooperation, clustering, and assortative mixing in dynamic networks.” *PNAS* 115(5): 951-956.

Nishi, A., H. Shirado, and N. Christakis. 2015. “Intermediate Levels of Network Fluidity Amplify Economic Growth and Mitigate Economic Inequality in Experimental Social Networks.” *Sociological Science* 2: 544-557.

Nishi, A., H. Shirado, D. Rand, and N. Christakis. 2015. “Inequality and visibility of wealth in experimental social networks.” *Nature* 526: 426-429.

Norton, M. and D. Ariely. 2011. “Building a Better America—One Wealth Quintile at a Time.” *Perspectives on Psychological Science* 6(1): 9-12.

Norton, M. 2014. “Unequality: Who Gets What and Why It Matters.” *Policy Insights from the Behavioral and Brain Sciences* 1(1): 151-155.

Piketty, T. and E. Saez. 2006. “The Evolution of Top Incomes: A Historical and International Perspective.” *AEA Papers and Proceedings* 96(2): 200-205.

Rand, D., M. Nowak, J. Fowler, and N. Christakis. 2014. “Static network structure can stabilize human cooperation.” *PNAS* 111(48): 17093-17098.

Shirado, H., F. Fu, J. Fowler, and N. Christakis. 2013. “Quality versus quantity of social ties in experimental cooperative networks.” *Nature Communications* 4: 2814.

Starmans, C., M. Sheskin, and P. Bloom. 2017. “Why people prefer unequal societies.” *Nature Human Behaviour* 1: 0082.

Tyler, T. “Procedural Justice Shapes Evaluations of Income Inequality: Commentary on Norton and Ariely (2011).” *Perspectives on Psychological Science* 6: 15-16.

Wang, J., S. Suri, and D. Watts. 2012. “Cooperation and assortativity with dynamic partner updating.” *PNAS* 109(36): 14363-14368.

Xiao, E. and C. Bicchieri. 2010. “When equality trumps reciprocity.” *Journal of Economic Psychology* 31(3): 456-470.

Figures

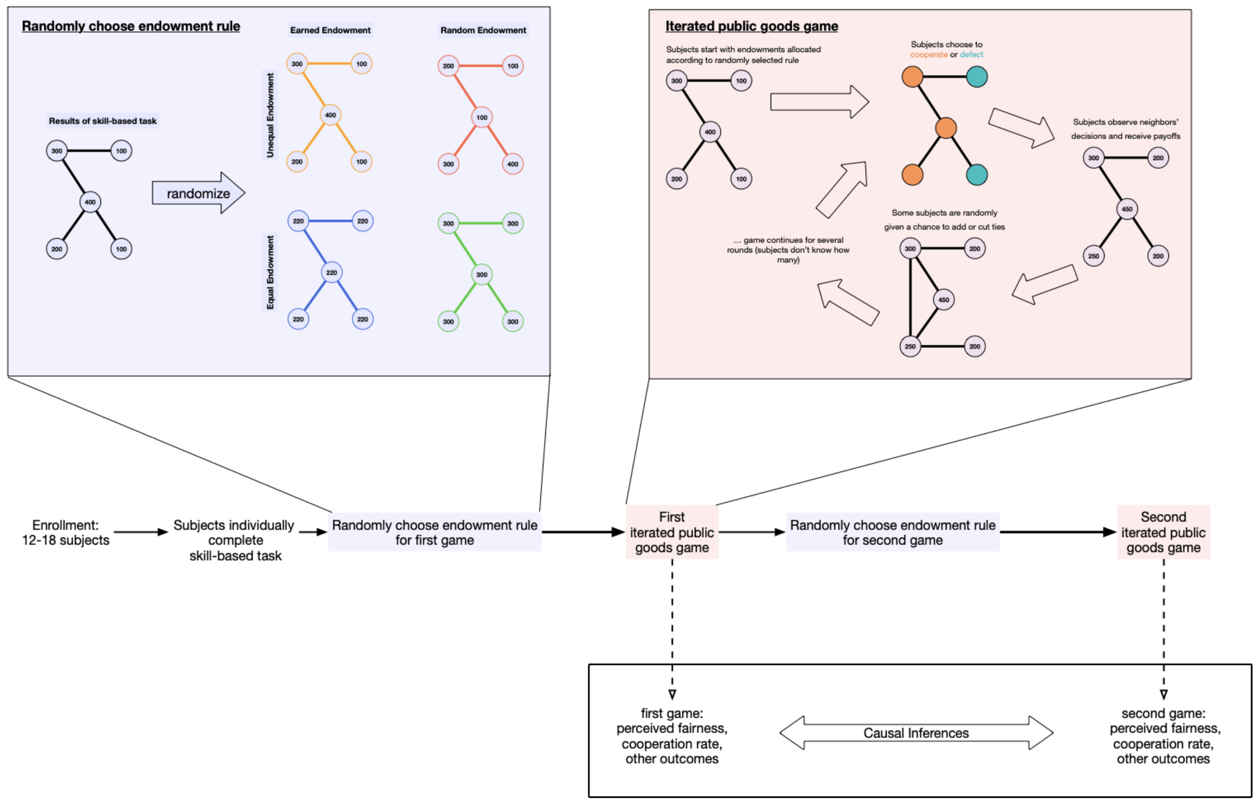


Figure 1 | Study design. Study flow for a single iteration of the experiment. The upper-left cell describes the four possible endowment regimes. The upper-right cell describes the progression of a standard public goods game with multiple rounds. The lower-right cell emphasizes the within-subjects design of our experiment.

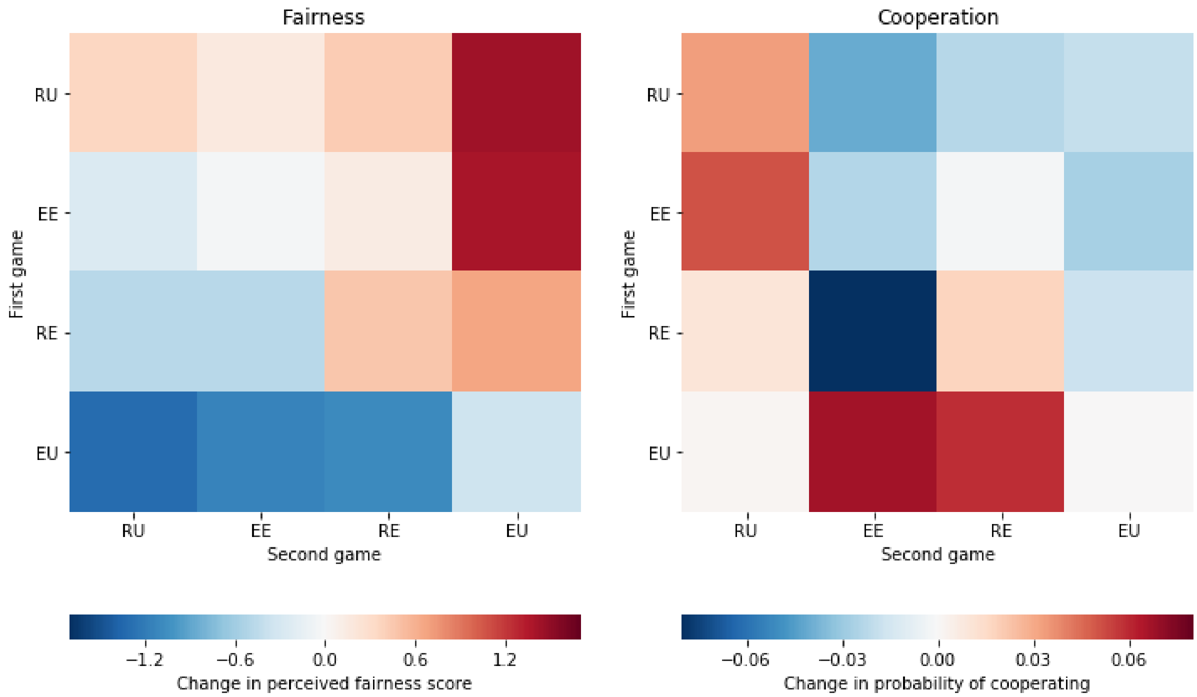


Figure 2 | Changes in outcomes as a function of changes in endowment regimes. The columns denote the endowment regime in the first public goods game, while the rows denote the endowment regime in the second public goods game. Red corresponds to positive average changes, while blue corresponds to negative average changes. EU: earned unequal; RU: random unequal; EE: earned equal; RE: random equal.

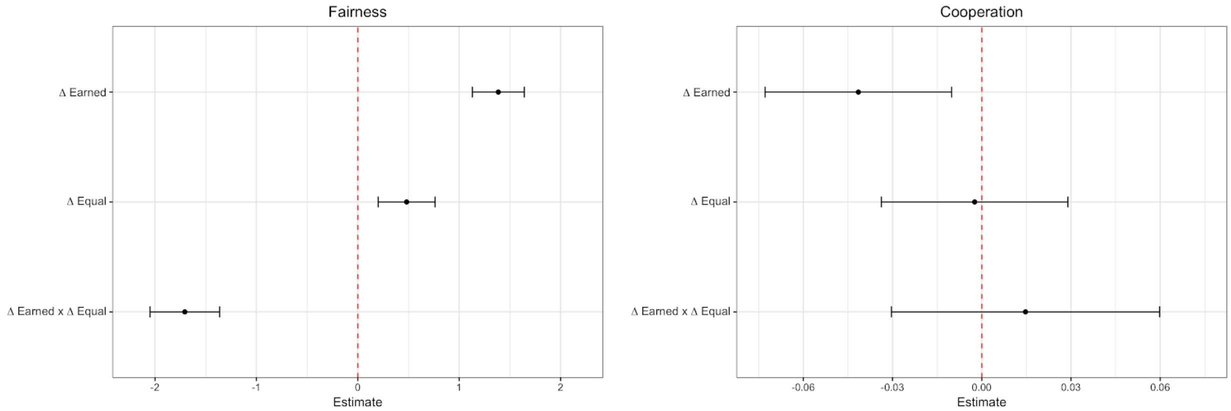


Figure 3 | Coefficient estimates with 95% confidence intervals. The point estimates correspond to the numbers mentioned in the text. The bars denote 95% confidence intervals constructed using clustered standard errors. The dashed red line at 0 corresponds to a null effect and is included to show which estimates are statistically significant.

Appendix 1: Experiment texts and figures

Screening

Welcome to our task!

Click 'Next' to continue.

Before we begin...

Please answer the following question.

[number 1] plus [number 2] equals…

[ANSWER HERE]

Word game tutorial

Tutorial (1/9)

In this task, you will play **three games**.

Your final payment for this task will be determined **at the end of the third game**.

Note that **only those players who finish the entirety of the task and click 'Submit HIT' at the end will receive any form of payment.**

Note also that once the first game starts, **if you remain idle for more than 20 seconds, you will be dropped from this HIT and hence will be ineligible for any form of payment.**

Click 'Next' to continue.

Tutorial (2/9)

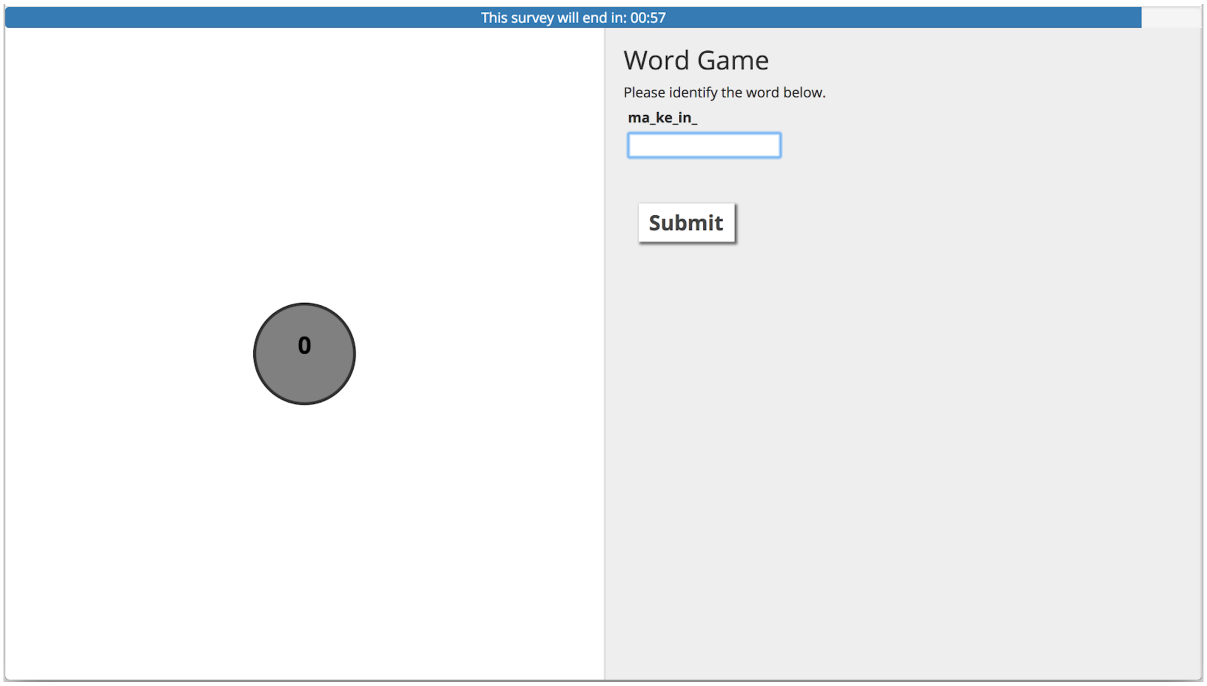
In the **first game**, you will be presented with a series of words with missing letters and asked to identify each word.

You start the **first game** with **0 points**, and each correct answer will add **100 points** to your score.

Click 'Next' to continue.

Tutorial (3/9)

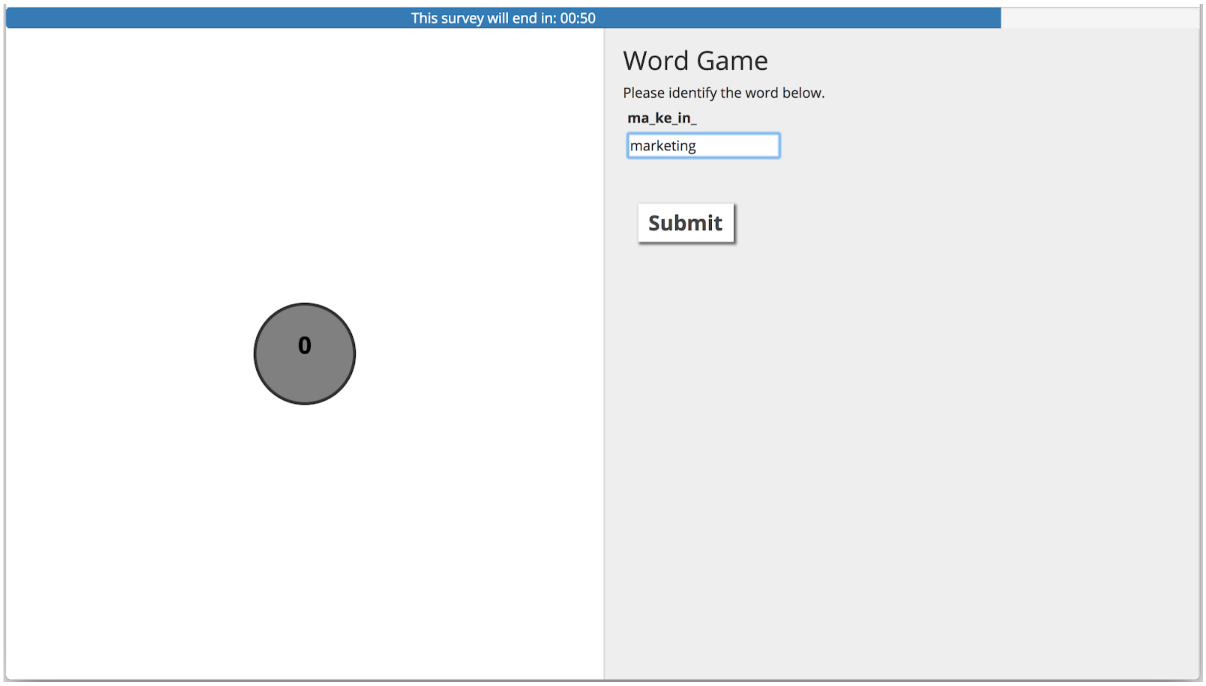
For example, if the first word that you are presented with is **ma_ke_in_**, then you will see the following screen:



Click 'Next' to continue.

Tutorial (4/9)

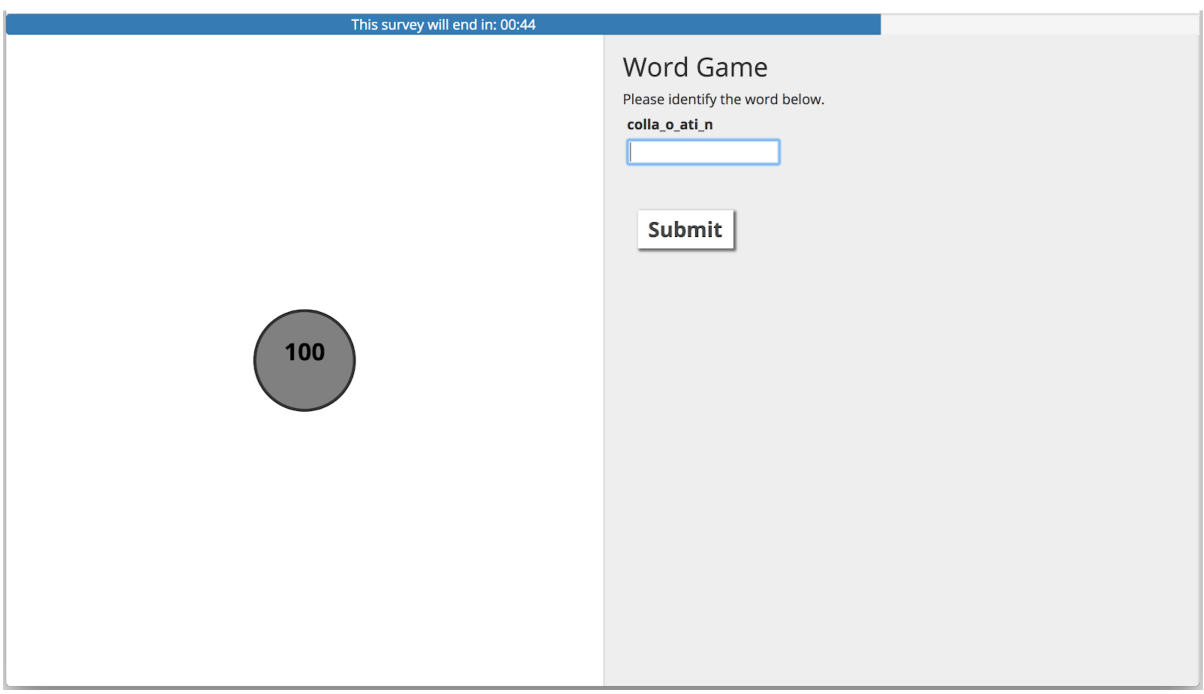
The correct answer, in this case, would be **marketing**, which you can enter in the box provided:



Click 'Next' to continue.

Tutorial (5/9)

Once you click 'Submit' or hit 'Enter' with the correct answer inside the box, **100 points** will be added to you, and the next word will be shown:



Click 'Next' to continue.

Tutorial (6/9)

Please click 'Submit' or hit 'Enter' after you enter each word to see the next word.

If you do not know a word, you can either guess or submit a blank answer.

Click 'Next' to continue.

Tutorial (7/9)

You have **60 seconds** to identify as many words as you can.

You will be able to see how many seconds you have left at the top.

There are a lot more words than you can identify in the given time frame, so you should not feel bad that you will not be able to identify all of them.

Click 'Next' to continue.

Tutorial (8/9)

Please enter all letters lowercase.

There may be some **plurals** among the words (e.g., chairs).

There may be some **proper nouns** among the words (e.g., texas).

Please make sure to enter all letters **lowercase** even if the word is a proper noun.

Click 'Next' to continue.

Tutorial (9/9)

Before you play the real game, we will let you play a **demo run**.

You will now start the demo run.

Note that when the run starts, you may be linked to other players in the game. Regardless of whether or not you are linked to others, remember that you are always the larger circle at the center.

Click 'Next' to continue.

First public goods game tutorial

Tutorial (1/10)

ONE OF THE FOLLOWING TREATMENT TEXTS (EE, EU, RE, RU)

> You will start the second game with a score that is the average of the scores of all players in the word game. This score could be lower than, equal to, or higher than your actual score from the word game depending on how your personal performance compares to how well the other players did in the task. For example, even if you performed well in the word game, if other players did not perform as well as you did, your score will unfortunately go down. Similarly, even if you performed poorly in the word game, if other players performed better than you did on average, then your score will go up.
>
> You will start the second game with your score from the word game. Other players similarly start this game with whatever score they were able to achieve in the word game. In other words, those who performed well in the word game start the second game with a higher score than others who did not perform as well.
>
> We will disregard your score from the word game. Instead, you will start the second game with a score that we randomly assign to you. All participants in your group are assigned the exact same score.
>
> We will disregard your score from the word game. Instead, you will start the second game with a score that we randomly assign to you. This score could be lower than, equal to, or higher than your actual score from the word game. In other words, even if you performed well in the word game, you could unfortunately still get a score that is much lower. Similarly, even if you performed poorly in the word game, you could still get a score that is much higher. It is highly likely that different players will be assigned different random scores.

Your score at the end of the word game was [SCORE IN WORD GAME].

The score you will start the second game with is [DEPENDS ON CONDITION].

The higher your score in the **second game** the higher your bonus payment will be at the end.

Please tell us how fair or unfair you find this rule for allocating scores.

Unfair

1

2

3

4

5

6

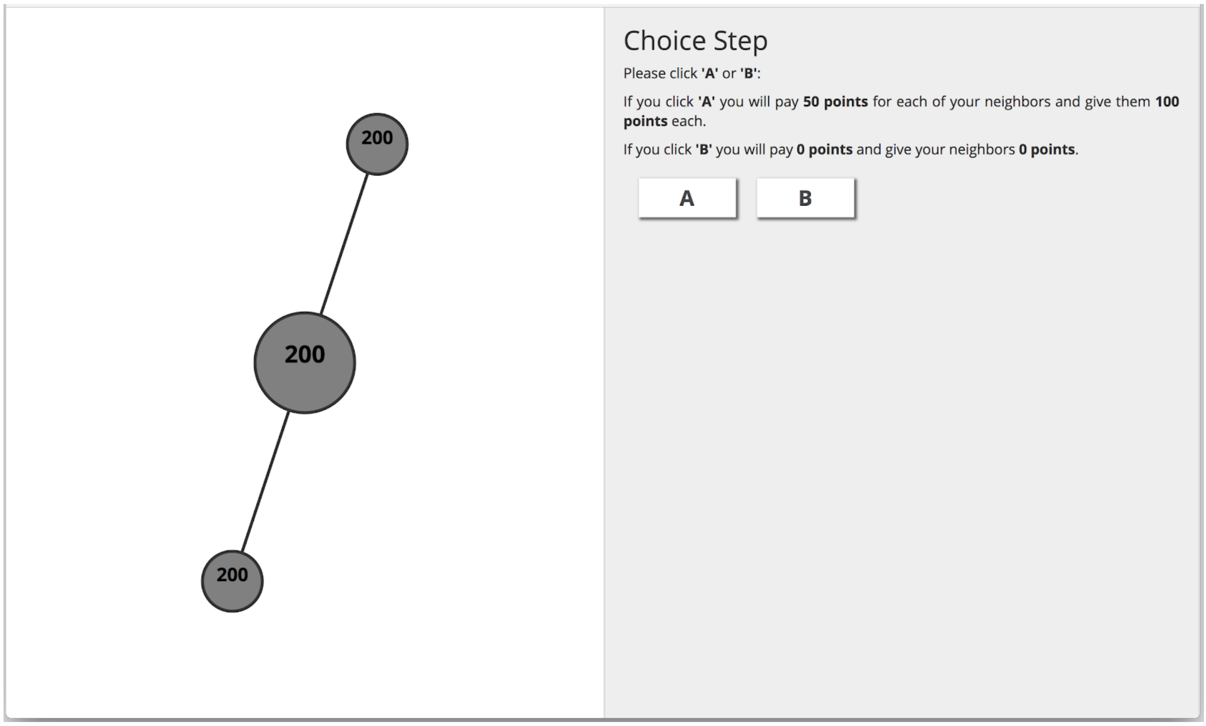
7

Fair

Tutorial (2/10)

In the **second game**, you will be linked together with other players to play a game where you can decide how much to contribute to one another.

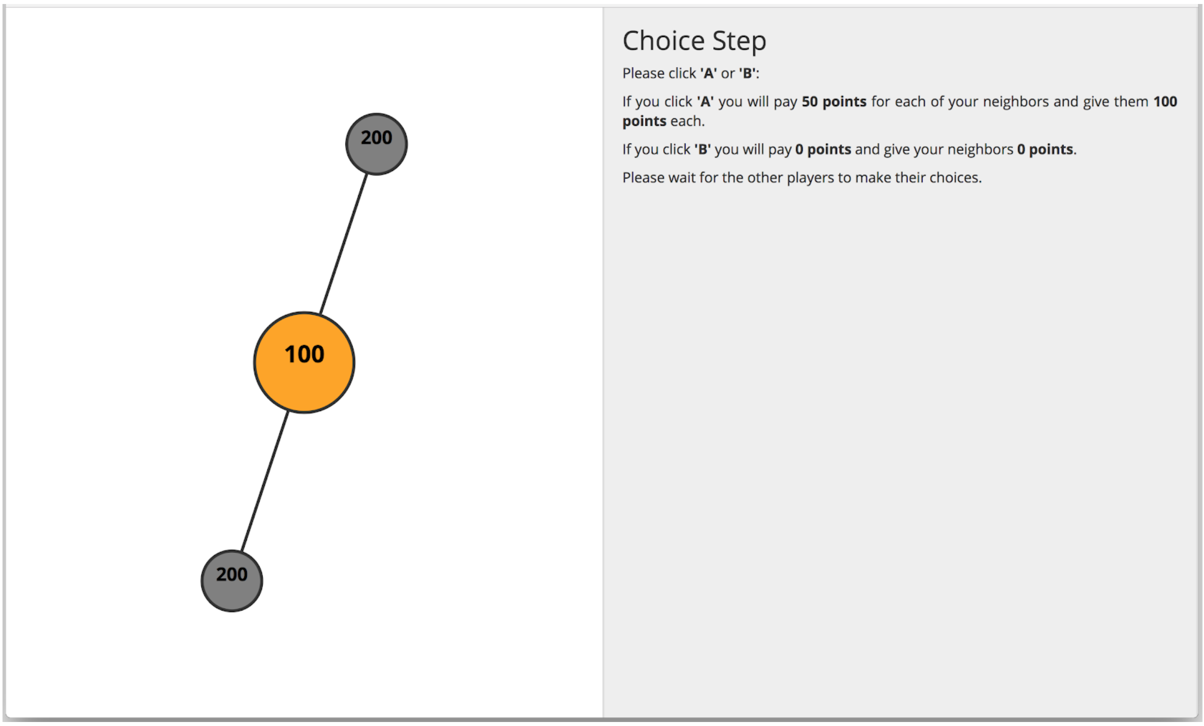
For example, the following screen shows a player with a score of 200 and who is connected to two neighbors also with scores 200 (the larger circle at the center is you):



Click 'Next' to continue.

Tutorial (3/10)

If you click **'A'** you will pay **50 points** for each player you are connected with, and we will give **100 points** to each player you are connected with:



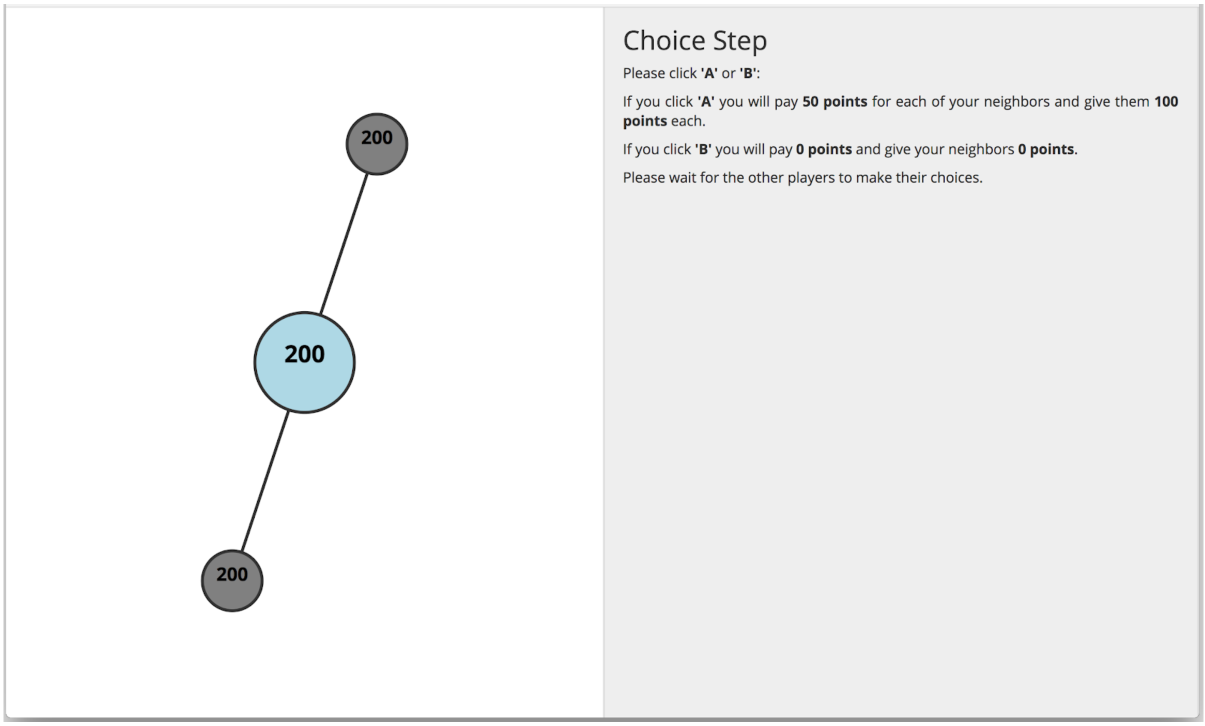
Note that you will not be able to see your neighbors' new scores and choices until after all players made their choices for that round.

Note also that your new score at the end of the round may end up being higher than what you are seeing at this stage (100 in this case) depending on how many of your neighbors also choose **'A'**.

Click 'Next' to continue.

Tutorial (4/10)

If you click **'B'** instead you will pay **0 points** and give **0 points** to each player you are connected with:



Once again, new scores and choices will be shown once all players make their choices.

Click 'Next' to continue.

Tutorial (5/10)

It is important to note that **if everybody chooses 'A'**, then everybody is guaranteed to be better off at the end of the game.

However, **if you choose 'A' and others choose 'B'**, then others will be better off, while you will be worse off.

Similarly, **if you choose 'B' and others 'A'**, then you will be better off, and others will be worse off.

You will be playing **multiple rounds** of this game.

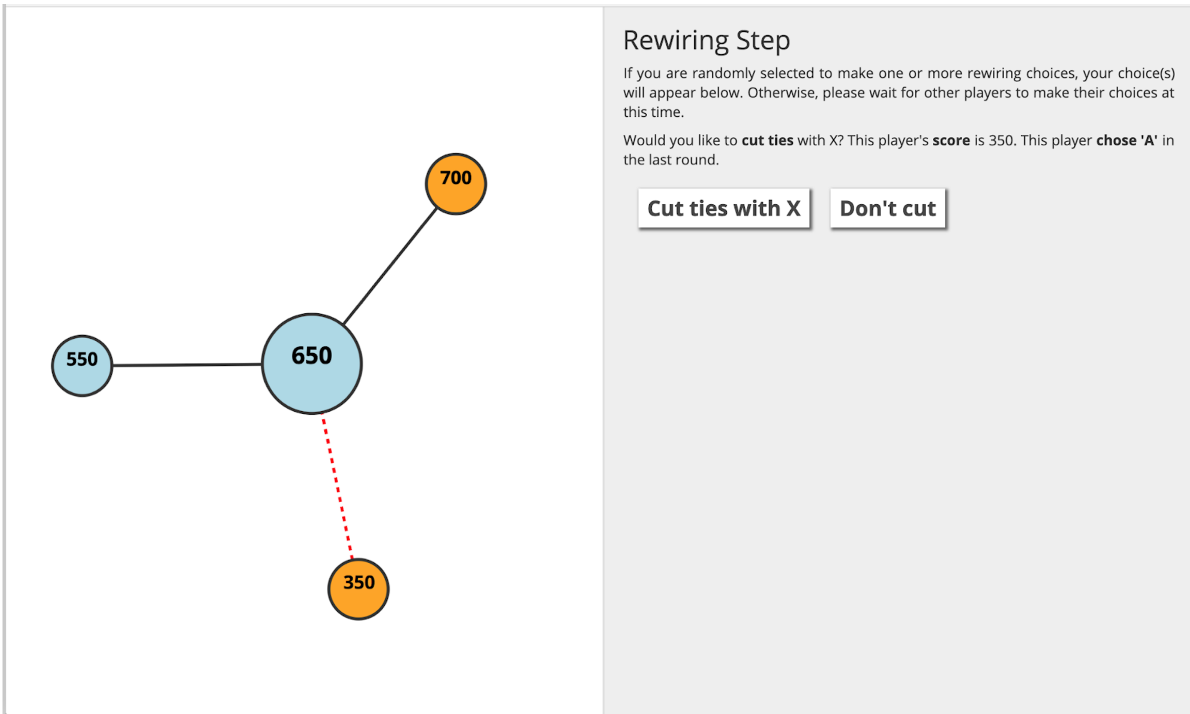
Click 'Next' to continue.

Tutorial (6/10)

After each round, a certain fraction of players will be randomly selected and allowed to cut or add ties with other players.

A tie can be cut from input from a single player.

For example, the following screen shows a player with a score of 650 and who is connected to three neighbors, with scores 350, 550, and 700. The player is asked whether to cut the tie to the player with score 350.

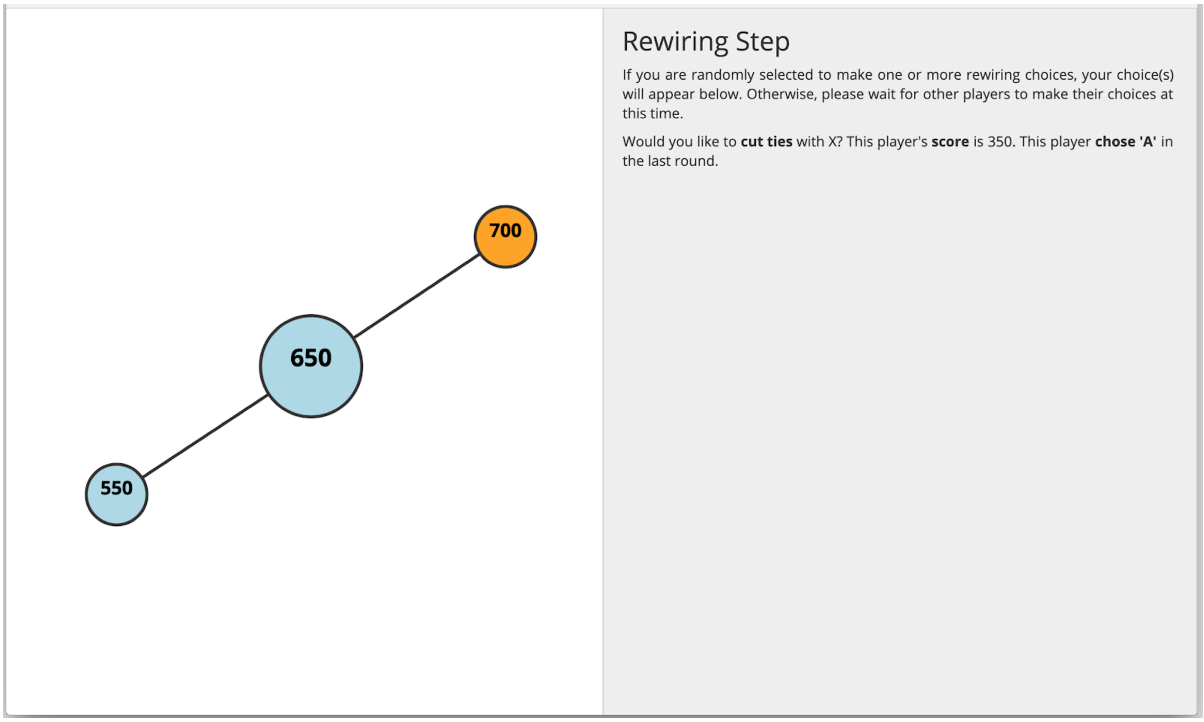


If you click 'Don't cut', then the tie will not be cut.

Click 'Next' to continue.

Tutorial (7/10)

However, if you click 'Cut ties with X', then your tie to player X will be cut at the end of the step:

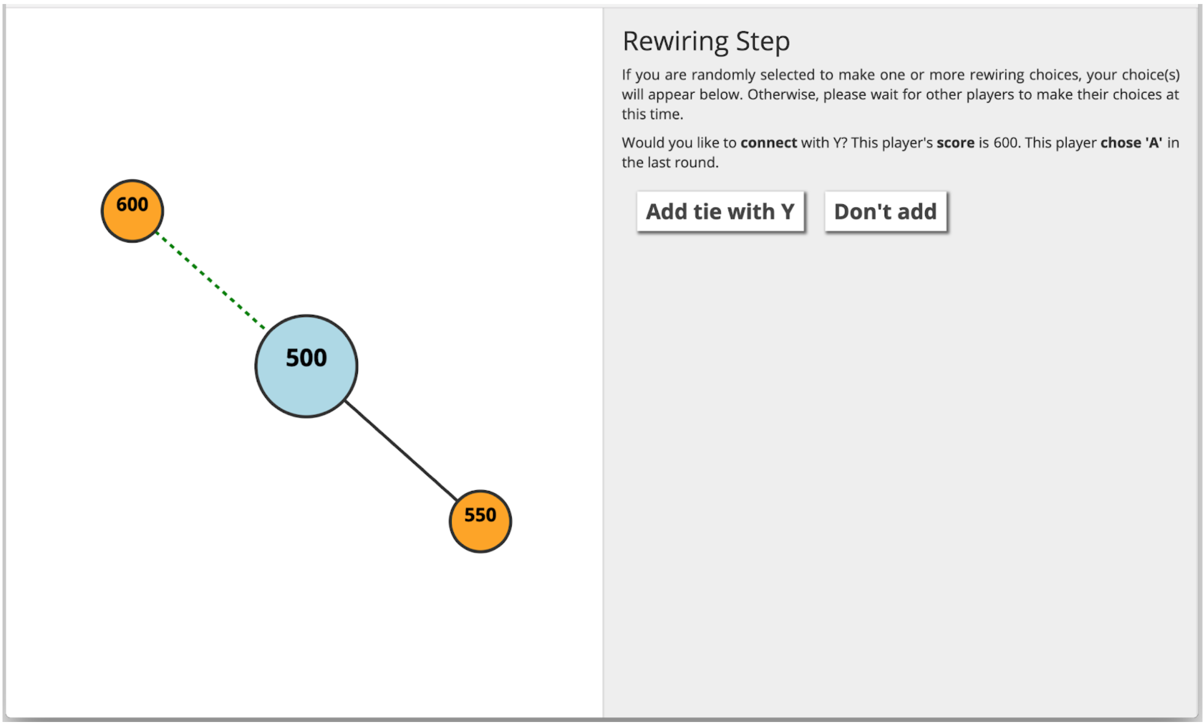


Click 'Next' to continue.

Tutorial (8/10)

A new tie will only be added if both players choose to connect.

For example, the following screen shows a player with a score of 500 and who is connected to one neighbor with a score of 550. The player is asked whether to add a new tie to the player who has a score of 600.



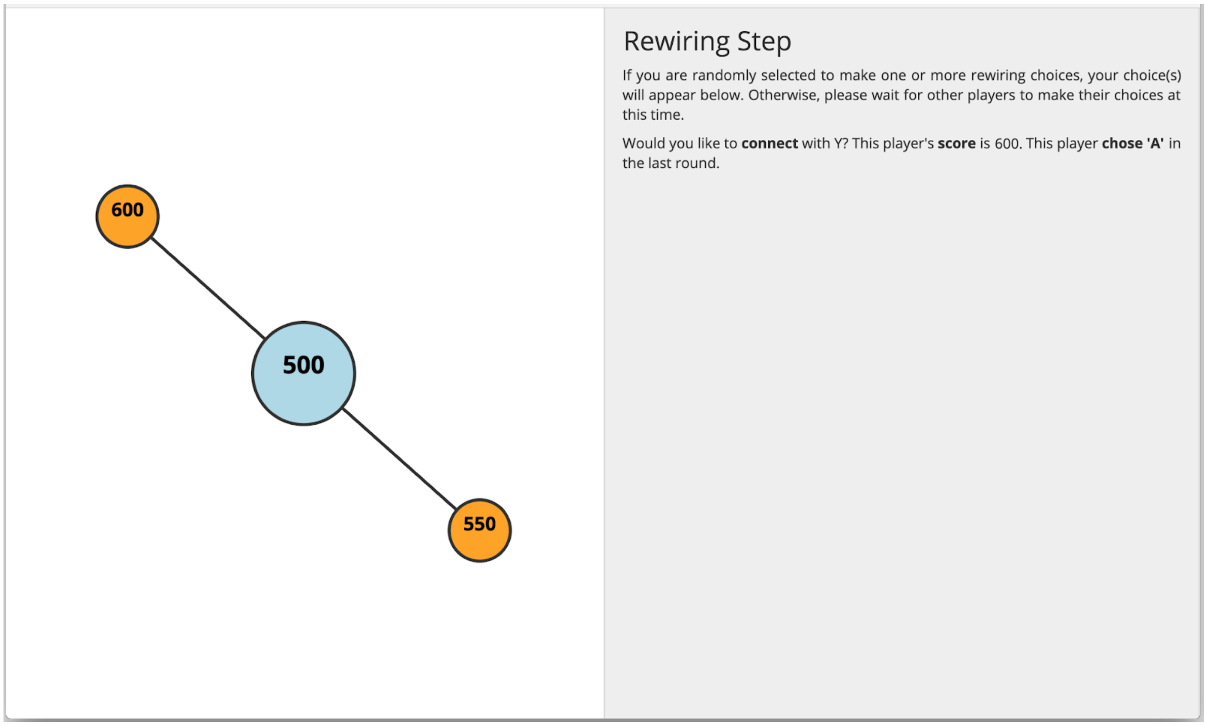
If you click 'Don't add', then the proposed tie will not be added.

Similarly, if you click 'Add tie with Y' but player Y clicks 'Don't add' on their end, then the proposed tie will still not be added.

Click 'Next' to continue.

Tutorial (9/10)

However, if player Y also clicks 'Add tie with X', then you will be connected to player Y:



Click 'Next' to continue.

Tutorial (10/10)

Before you play the real second game, we will let you play a **demo run**.

You will now start the demo run.

Click 'Next' to continue.

Second public goods game tutorial

Tutorial (1/4)

We recorded the score you reached at the end of the second game.

You will now play the third and final game of this task.

Click 'Next' to continue.

Tutorial (2/4)

The rules of this game are identical to the second game you just played.

IF CONDITION DIFFERENT FROM BEFORE

However, two things have changed:

- We will randomly choose a new set of connections to start the game.
- Your starting points will be chosen differently this time.

IF CONDITION SAME AS BEFORE

We will randomly choose a new set of connections to start the game.

Click 'Next' to continue.

Tutorial (3/4)

ONE OF THE FOLLOWING TREATMENT TEXTS (EE, EU, RE, RU)

You will start the third game with a score that is the average of the scores of all players in the word game. This score could be lower than, equal to, or higher than your actual score from the word game depending on how your personal performance compares to how well the other players did in the task. For example, even if you performed well in the word game, if other players did not perform as well as you did, your score will unfortunately go down. Similarly, even if you performed poorly in the word game, if other players performed better than you did on average, then your score will go up.

You will start the third game with your score from the word game. Other players similarly start this game with whatever score they were able to achieve in the word game. In other words, those who performed well in the word game start the third game with a higher score than others who did not perform as well.

We will disregard your score from the word game. Instead, you will start the third game with a score that we randomly assign to you. All participants in your group are assigned the exact same score.

We will disregard your score from the word game. Instead, you will start the third game with a score that we randomly assign to you. This score could be lower than, equal to, or higher than your actual score from the word game. In other words, even if you performed well in the word game, you could unfortunately still get a score that is much lower. Similarly, even if you performed poorly in the word game, you could still get a score that is much higher. It is highly likely that different players will be assigned different random scores.

Your score at the end of the word game was [SCORE IN WORD GAME].

The score you will start the third game with is [DEPENDS ON CONDITION].

The higher your score in the **third game** the higher your bonus payment will be at the end.

Please tell us how fair or unfair you find this rule for allocating scores.

Unfair

1

2

3

4

5

6

7

Fair

Tutorial (4/4)

Your performance bonus will be based on **your final scores at the end of the second and third games**.

You will now start the third game.

Remember that you will be playing MULTIPLE rounds of this game.

Click 'Begin' to join the game.

After you click 'Begin', please stay on this page as you may be dropped for being idle if you don't make your next move within **<u>20 seconds</u>** when it appears.

<u>Trust survey shown at the end of each public goods game</u>

Survey Step

Do you think that most of your neighbors tried to take advantage of you when they got the chance, or did they try to be fair?

Most of them tried to take advantage of me.

1

2

3

4

5

6

7

Most of them tried to be fair.

<u>End survey</u>

IF THE TWO CONDITIONS WERE DIFFERENT

Survey Step

Remember that you played the community game **twice**.

Please compare the **<u>first version</u>** of the game with the **<u>second version</u>** and tell us which version seems **<u>more fair</u>** to you.

Both versions are described below.

The rules for the first version of the game:

[TREATMENT TEXT 1]

The rules for the second version of the game:

[TREATMENT TEXT 2]

Please tell us which version seems MORE FAIR to you.

first version

second version

If you had the chance to play this game one more time, which version would you like to play?

first version

second version

Survey Step

Recall the many choices you had to make whether to **take action A (and contribute points to your neighbors)** or **take action B (and NOT contribute points to your neighbors)** during the game.

Which of the following factors were most influential in making you choose action A (and contribute points to your neighbors)? You can check one or more boxes.

I wanted other players to increase their scores.

I wanted to encourage other players to choose A too.

Most of my neighbors chose A in the previous round.

Most of my neighbors chose B in the previous round.

Most of my neighbors had similar scores compared to me.

Most of my neighbors had higher scores compared to me.

Most of my neighbors had lower scores compared to me.

I found the rule of initial score allocation to be fair.

I didn't find the rule of initial score allocation to be fair.

Other factor not listed here (type your reason below).

Survey Step

How many other HITs have you participated in that required you to interact with other players like this HIT?

[ANSWER HERE]

How old are you?

[ANSWER HERE]

What gender do you identify with?

Male

Female

Other

What race/ethnicity do you identify with?

White

Black

Hispanic

Asian

Other

What is your level of education?

Less than high school diploma

High school diploma or equivalent

Some college

College degree

Graduate degree

Other

What is your yearly income in US dollars?

Less than $20,000

$20,000 to $39,999

$40,000 to $59,999

$60,000 to $79,999

$80,000 to $99,999

More than $100,000

Which of the following best describes your political orientation?

Very liberal

Liberal

Middle of the road

Conservative

Very conservative

Are you located in the U.S.?

Yes

No

Appendix 2: Power calculations

To determine the target sample size we need for this experiment, we conducted a power study. Following Snijders (2005), we base our analysis on the approximate relationship for a two-sided test

$$\frac{\gamma}{\text{s.e.}(\widehat{\gamma})} \approx z_{1-(\alpha/2)} + z_{1-\beta} = K_z,$$

where γ is the true multilevel model coefficient on a treatment effect, $\text{s.e.}(\widehat{\gamma})$ is the standard error of the estimated treatment effect, $\alpha = 0.05$ is the significance level, $1-\beta = 0.8$ is the power, and $K_z = \Phi^{-1}\left(1-\frac{\alpha}{2}\right) + \Phi^{-1}(1-\beta)$ is called the non-centrality parameter. Squaring this relationship, we have

$$\frac{\gamma^2}{\text{var}(\widehat{\gamma})} \approx K_z^2.$$

To estimate sample size, we need an expression for the variance of γ. Following Moerbeek, Breukelen, and Berger (2000), we estimate the expected variance of our multilevel regression coefficient on a treatment variable (assuming 0/1 coding) as:

$$\text{var}(\widehat{\gamma}) = \frac{1}{n_3} K_\sigma,$$

where K_σ is defined as

$$K_\sigma = 4 \times \frac{\left[n_1 n_2 \sigma_3^2 + n_1 \sigma_2^2 + \sigma_1^2\right]}{n_1 n_2} = 4 \times \left[\sigma_3^2 + \frac{\sigma_2^2}{n_2} + \frac{\sigma_1^2}{n_1 n_2}\right].$$

n_i and σ_i^2 are the sample size and the variance of the dependent variable for level i, where $i = 1, 2, 3$ for rounds, subjects, and networks, respectively. To find an expression for the number of networks, n_3, given all of the other parameter values, we start from the equation above, plugging the expression for $\text{var}(\widehat{\gamma})$ in and rearranging to solve for n_3:

$$\gamma^2 \frac{n_3}{K_\sigma} \approx K_z^2 \Longleftrightarrow n_3 \approx \frac{K_z^2 K_\sigma}{\gamma^2}.$$

We set $\gamma = 0.06$, $\alpha = 0.05$, and $1-\beta = 0.8$. We use a multilevel model estimated on the replication data published by Nishi et al (2015) to approximate $\sigma_1^2 = 0.14$, $\sigma_2^2 = 0.08$, and $\sigma_3^2 = 0.02$. Finally, we assume there will be $n_1 = 10$ rounds per game and $n_2 = 12$ subjects per network. Solving for n_3, we obtain a preliminary desired sample size of $n_3^b = 243$ networks for a between-subjects design.

Finally, to account for our within-subjects design, we use the approximate relationship $n_3^w = n_3^b(1-\rho)/2$, where ρ is the within-subjects correlation in the outcome across games, and n_3^w is the within-subjects target sample size (Maxwell and Delaney 2004). Taking $\rho = 0$ to be conservative, we obtain a target within-subjects sample size of $n_3 = 122$ networks.

In practice, we ended up collecting more data on 1870 subjects nested inside 160 networks (Table A2.1). (For reference, each network session took about 25-30 minutes to complete, and subjects were paid an average of $4 for their participation.)

Table A2.1. Number of subjects by condition (numbers inside parentheses complete samples).

1st public goods game		2nd public goods game		
Earned	Equal	Earned	Equal	Number of subjects
0	0	0	0	114 (103)
			1	124 (117)
		1	0	127 (119)
			1	105 (99)
	1	0	0	111 (108)
			1	129 (120)
		1	0	123 (118)
			1	122 (113)
1	0	0	0	103 (91)
			1	106 (100)
		1	0	128 (118)
			1	112 (108)
	1	0	0	111 (105)
			1	113 (109)
		1	0	132 (129)
			1	110 (102)
				1870 (1759)

Citations

Maxwell, S., and H. Delaney. 2004. *Designing experiments and analyzing data: a model comparison perspective*, 2nd edition. New Jersey: Lawrence Erlbaum Associates.

Moerbeek, M., G. Breukelen, and M. Berger. 2000. "Design Issues for Experiments in Multilevel Populations." *Journal of Educational and Behavioral Statistics* 25(3): 271-284.

Nishi, A., H. Shirado, D. Rand, and N. Christakis. 2015. "Inequality and visibility of wealth in experimental social networks." *Nature* 526: 426-429.

Snijders, T. “Power and Sample Size in Multilevel Linear Models.” *Encyclopedia of Statistics in Behavioral Science, Volume 3*. Ed. B. Everitt, Ed. D. Howell. Chichester: Wiley, 2005. 1570-1573.

Appendix 3: Variables and demographics

The final dataset has a large number of variables. The first set of variables are those that help us precisely identify each unique network session, subject, public goods game, round, and action as well as the endowment regimes associated with them. These include: (i) network id, (ii) subject id, (iii) a dummy variable determining whether an observation comes from the first or the second public goods game, (iv) whether initial endowment was earned, (v) whether initial endowment was equal, and (vi) round. There are also variables for the exact (vii) date and (viii) start time of each session.

The second set of variables relate to the skill-based task and ego's fairness perceptions of the subsequent initial endowment allocation mechanism, including: (i) ego's score in the skill-based task, (ii) the change in ego's score from the skill-based task to the public goods game, (iii) fairness score ego assigned to the endowment allocation mechanism of the first public goods game, (iv) fairness score ego assigned to the endowment allocation mechanism of the second public goods game, (v) which game ego considers to be more fair (asked at the end of the experiment), and (vi) which game ego would prefer to play again if he/she had the chance (asked at the end of the experiment). Subjects were also asked to answer a survey question regarding (vii) the reason(s) why they chose to cooperate (asked at the end of the experiment).

The third set of variables include those that relate to subjects' decisions in the two public goods games. These are: (i) ego's cooperation decision in a given round, (ii) ego's score at the time of cooperation decision, (iii) ego's tie formation decision in a given round, (iv) ego's score at the time of tie formation decision, (v) alter's score at the time of tie formation decision, (vi) ego's cooperation decision prior to tie formation decision, (vii) alter's cooperation decision prior to tie formation decision, (viii) ego's tie breakage decision in a given round, (ix) ego's score at the time of tie breakage decision, (x) alter's score at the time of tie breakage decision, (xi) ego's cooperation decision prior to tie breakage decision, and (xii) alter's cooperation decision prior to tie breakage decision. In addition, there are variables for (xiii) ego's score at the end of the first public goods game, (xiv) ego's score at the end of the second public goods game, (xv) trust score ego assigned regarding other subjects' behavior at the end of the first public goods game, (xvi) trust score ego assigned regarding other subjects' behavior at the end of the second public goods game, and (xvii) ego's connections in a given round.

Next is the set of variables related to the overall network. These include: (i) average cooperation in a given round in the whole network, (ii) average cooperation in a given round among ego and his/her immediate connections only, (iii) gini index in a given round in the whole network, (iv) gini index in a given round among ego and his/her immediate connections only, and (v) number of subjects in a given round.

In addition to the variables described above, certain demographic information was also collected from the subjects, including: (i) the number of similar networked tasks they participated in before (i.e., experience), (ii) age, (iii) gender, (iv) race, (v) level of education, (vi) level of income, (vii) political orientation, and (viii) whether they are located in the US. An additional variable for (ix) country based on IP address was also generated. The overall demographic composition of the sample can be described as follows:

(i) Mean experience is 301, while the median is only 2. The huge discrepancy between the mean and the median here is caused by the fact that there are some subjects who stated that they played a really large number of similar games.

(ii) Mean age is 37, while the median is 34.

(iii) The sample is 47.2% male, 52.2% female, and 0.6% other.
(iv) The sample is 78.7% White, 7.0% Black, 7.1% Hispanic, 5.3% Asian, and 1.9% other.
(v) The level of education in the sample is 0.5% less than high school, 10.8% high school, 30.1% some college, 43.2% college, 15.3% graduate, and 0.1% other.
(vi) The level of income in the sample is 19.2% less than $20K, 26.6% between $20-40K, 23.9% between $40-60K, 14.7% between $60-80K, 7.6% between $80-100K, and 8.0% more than $100K.
(vii) The sample is 17.3% very liberal, 33.8% liberal, 27.6% middle of the road, 17.4% conservative, and 3.9% very conservative.

Appendix 4: Within-subjects results

Fairness perceptions

The effect of change in endowment regime on change in fairness perceptions can be estimated using the following within-subjects models.

$$\Delta fairness_{ij} = \delta_0 + \delta_1 \Delta earned_i + \delta_2 \Delta equal_i + \delta_3 \Delta earned_i \times equal_i + e_{ij} \quad \text{(A4.1)}$$

$$\Delta fairness_{ij} = \delta_0 + \delta_1 \Delta earned_i + \delta_2 \Delta equal_i + \delta_3 \Delta earned_i \times equal_i + u_i + e_{ij} \quad \text{(A4.2)}$$

Equation A4.1 (identical to Equation 1 in the main text) uses standard errors clustered at the network (i) level, while Equation A4.2 denotes a linear mixed-effects model, thus explicitly decomposing the error term into two parts. Table A4.1 presents results from these within-subjects models. The upper part of the table (see coefficient estimates with Δ before them) presents results based on the model that treats change in endowment regime as a continuous variable. The lower part of the table (see coefficient estimates with "Change to" before them) presents results based on the model that treats change in endowment regime as a categorical variable. See the main text for a discussion of these results. Estimates based on the model with clustered standard errors and the linear mixed-effects model are practically identical.

Cooperation patterns

The effect of change in endowment regime on change in cooperation decisions can be estimated using the following within-subjects models.

$$\Delta cooperation_{ijl} = \delta_0 + \delta_1 \Delta earned_i + \delta_2 \Delta equal_i + \delta_3 \Delta earned_i \times equal_i + \sum_{l=1}^{9} \delta_{4l}\, round_{il} + e_{ijl} \quad \text{(A4.3)}$$

$$\Delta cooperation_{ijl} = \delta_0 + \delta_1 \Delta earned_i + \delta_2 \Delta equal_i + \delta_3 \Delta earned_i \times equal_i + \sum_{l=1}^{9} \delta_{4l}\, round_{il} + u_i + u_{ij} + e_{ijl} \quad \text{(A4.4)}$$

Equation A4.3 (identical to Equation 2 in the main text) uses standard errors clustered at the network (i) level, while Equation A4.4 denotes a linear mixed-effects model, thus explicitly decomposing the error term into two parts. Table A4.2 presents results from these within-subjects models. Once again, the upper part of the table (see coefficient estimates with Δ before them) presents results based on the model that treats change in endowment regime as a continuous variable, while the lower part of the table (see coefficient estimates with "Change to" before them) presents results based on the model that treats change in endowment regime as a categorical variable. See the main text for a discussion of these results. Estimates based on the model with clustered standard errors and the linear mixed-effects model are practically identical. Finally, Table A4.3 presents results based on the instrumental variable model discussed in the main text, and Figure A4.1 visualizes the inverse relationship between fairness perceptions and cooperation.

The results are robust to (i) fitting the models on network- rather than individual-level data (Table A4.4), (ii) including additional predictors for change in subject score (from skill-based task to endowment allocation), subject score at the time of the cooperation decision, and experience (number of similar games subject participated in) (Table A4.5), and (iii) using an ordinal logistic model instead of OLS (Table A4.6).

Table A4.1. Fairness perceptions as a function of endowment regime (within-subjects).
Change in endowment regime treated as a continuous predictor

	Model with clustered standard errors	Linear mixed-effects model
Intercept	0.071 (0.060)	0.073 (0.060)
Δ Earned	1.386 (0.131)***	1.387 (0.120)***
Δ Equal	0.482 (0.143)**	0.481 (0.119)***
Δ Earned x Equal	-1.706 (0.175)***	-1.705 (0.168)***

Change in endowment regime treated as a categorical predictor

	Model with clustered standard errors	Linear mixed-effects model
Intercept	0.183 (0.119)	0.177 (0.151)
Change to Earned	0.393 (0.318)	0.437˙ (0.260)
Change to Random	-0.627* (0.263)	-0.612* (0.266)
Change to Equal	-0.427 (0.278)	-0.444˙ (0.261)
Change to Unequal	0.433 (0.280)	0.389 (0.259)
Change to Earned, Equal	0.096 (0.454)	0.058 (0.458)
Change to Earned, Unequal	-0.268 (0.465)	-0.268 (0.447)
Change to Random, Equal	-0.139 (0.418)	-0.147 (0.460)
Change to Random, Unequal	-0.165 (0.438)	-0.135 (0.457)

The number of samples (n) for the models fit is 1803, clustered inside 160 sessions. The numbers inside the parentheses are standard errors. Stars denote p-values: ˙ p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

Table A4.2. Cooperation decision as a function of endowment regime (within-subjects).
Change in endowment regime treated as a continuous predictor

	Models with clustered standard errors	Linear mixed-effects models
Intercept	-0.132 (0.012)***	-0.128 (0.013)***
Δ Earned	-0.042 (0.016)*	-0.041 (0.017)*
Δ Equal	-0.002 (0.016)	-0.004 (0.017)
Δ Earned x Equal	0.015 (0.023)	0.016 (0.024)

Change in endowment regime treated as a categorical predictor

	Models with clustered standard errors	Linear mixed-effects models
Intercept	-0.124*** (0.021)	-0.120*** (0.020)
Change to Earned	-0.060* (0.030)	-0.059* (0.029)
Change to Random	-0.009 (0.029)	-0.010 (0.030)
Change to Equal	0.013 (0.031)	0.007 (0.030)
Change to Unequal	-0.016 (0.029)	-0.016 (0.029)
Change to Earned, Equal	-0.003 (0.044)	0.004 (0.052)
Change to Earned, Unequal	0.052 (0.044)	0.048 (0.051)
Change to Random, Equal	0.046 (0.052)	0.055 (0.052)
Change to Random, Unequal	0.068 (0.056)	0.066 (0.052)

The number of samples (*n*) for the models fit is 16098, clustered inside 1803 players and 160 sessions. The numbers inside the parentheses are standard errors. Stars denote p-values: ˙ p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

Table A4.3. Instrumental variable model.

First stage (F-statistic = 66.77)

	Models with clustered standard errors	Linear mixed-effects models
Intercept	0.071 (0.060)	0.073 (0.060)
Δ Earned	1.386 (0.131)***	1.387 (0.120)***
Δ Equal	0.482 (0.143)**	0.481 (0.119)***
Δ Earned x Equal	-1.706 (0.175)***	-1.705 (0.168)***

Second stage

	Models with clustered standard errors	Linear mixed-effects models
Intercept	-0.130 (0.012)***	-0.127 (0.013)***
Predicted Δ Fairness	-0.022 (0.010)*	-0.023 (0.012)*

The number of samples (*n*) for the first-stage model fit is 1803, clustered inside 160 sessions. The number of samples (*n*) for the second-stage model fit is 16098, clustered inside 1803 players and 160 sessions. The numbers inside the parentheses are standard errors. Stars denote p-values: ˙ p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

Table A4.4. Estimates from network-level models.

	Fairness	Cooperation
Intercept	0.062 (0.060)	-0.126 (0.013)***

Δ Earned	1.387 (0.120)***	-0.038 (0.016)*
Δ Equal	0.469 (0.120)***	-0.003 (0.016)
Δ Earned x Equal	-1.690 (0.169)***	0.013 (0.024)

The number of samples (*n*) for the fairness model is 160 sessions. The number of samples (*n*) for the cooperation model is 1440 session-rounds, clustered inside 160 sessions. The numbers inside the parentheses are standard errors. Stars denote p-values: ˙ p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

Table A4.5. Estimates from models that control for additional predictors.

	Fairness	Cooperation
Intercept	0.099 (0.054)˙	-0.132 (0.013)***
Δ Earned	1.287 (0.126)***	-0.042 (0.017)*
Δ Equal	0.417 (0.121)**	0.001 (0.017)
Δ Earned x Equal	-1.455 (0.152)***	0.009 (0.025)
Δ Change in score after task	0.002 (0.000)***	
Δ Score before decision		-0.000 (0.000)***
Experience		-0.000 (0.000)

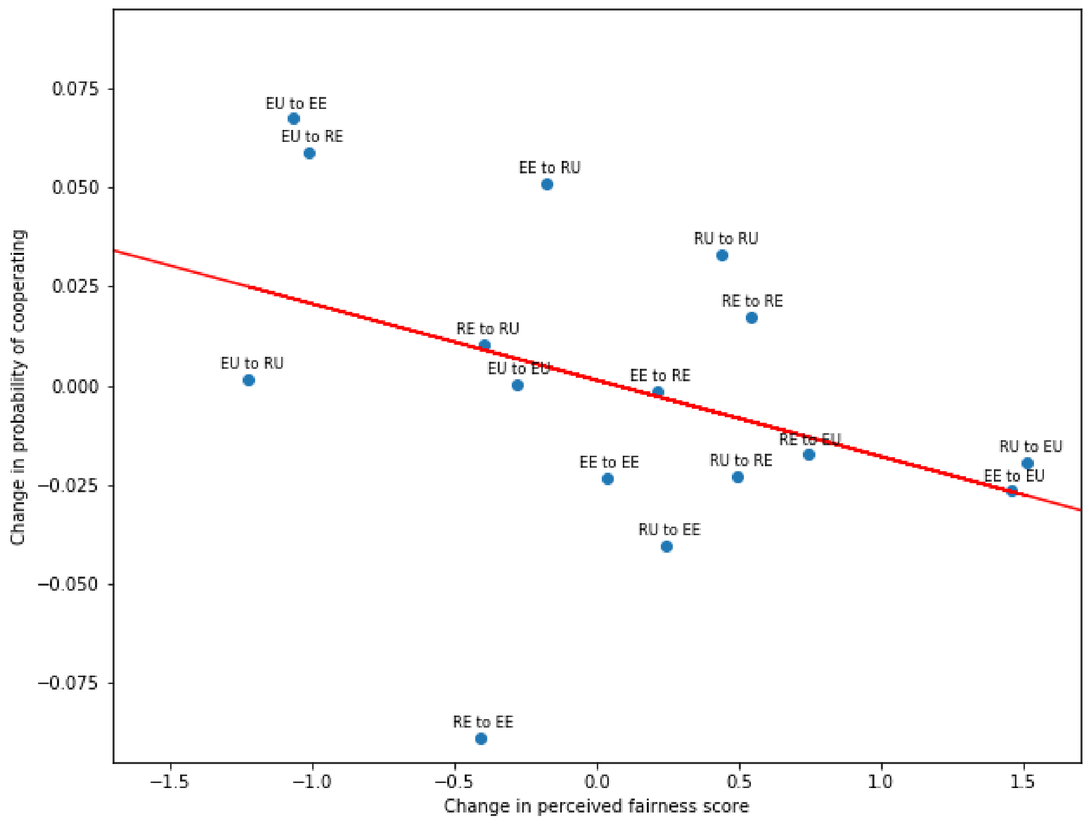
The number of samples (*n*) for the fairness model is 1803, clustered inside 160 sessions. The number of samples (*n*) for the cooperation model is 16098, clustered inside 1803 players and 160 sessions. The numbers inside the parentheses are standard errors. Stars denote p-values: ˙ p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

Table A4.6. Estimates from ordinal logistic regression models.

	Fairness	Cooperation
Δ Earned	1.099 (0.098)***	-0.221 (0.095)*
Δ Equal	0.381 (0.093)***	-0.013 (0.094)
Δ Earned x Equal	-1.304 (0.134)***	0.083 (0.133)

The number of samples (*n*) for the fairness model is 1803, clustered inside 160 sessions. The number of samples (*n*) for the cooperation model is 16098, clustered inside 1803 players and 160 sessions. The numbers inside the parentheses are standard errors. Stars denote p-values: ˙ p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

Figure A4.1. Change in perceived fairness score vs change in probability of cooperating.


EU to EE
EU to RE
EE to RU
RU to RU
RE to RE
RE to RU
EU to RU
EU to EU
EE to RE
RE to EU
RU to EU
EE to EU
EE to EE
RU to RE
RU to EE
RE to EE
0.075
0.050
0.025
0.000
−0.025
−0.050
−0.075
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
Change in probability of cooperating
Change in perceived fairness score

Appendix 5: Between-subjects results

The main text does not present between-subjects results for the following two reasons. First, we powered our study with the within-subjects design in mind, so the study does not have enough power to detect small between-subjects differences (see Appendix 2). Second, the within-subjects models are better-suited for making causal inferences, because patterns based on between-subjects comparisons could be masking heterogenous individual-level results. We present a between-subjects analysis in this appendix because many other studies have used a between-subjects design (e.g., Wang, Suri, and Watts 2012, Shirado et al 2013, Rand et al 2014, Nishi et al 2015, Nishi, Shirado, and Christakis 2015), and so we expect that some researchers may be interested in an analysis that is more easily compared to previous studies.

Fairness perceptions

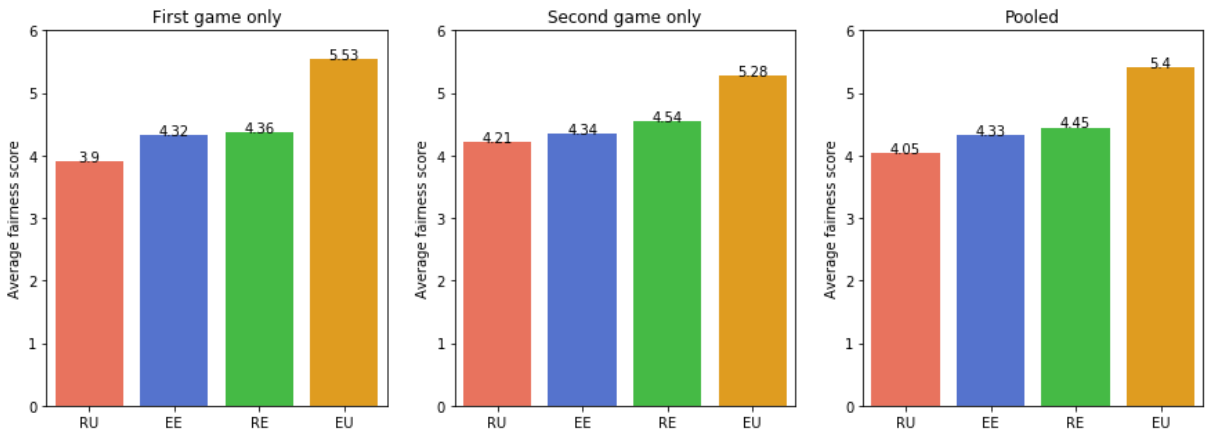
Figure A5.1 below presents average fairness scores by endowment regime for (i) the first game only, (ii) the second game only, and (iii) both games pooled together. As can be seen in this figure, regardless of which specific plot we focus on, the fairness ordering of the four regimes is always, from least to most fair: random unequal (RU, least fair), earned equal (EE), random equal (RE), and earned unequal (EU, most fair). EU is more fair compared to the other three conditions by a large margin (1 point or more on a scale of 1 to 7). The fairness scores of the two equal regimes are closer to one another (EE 4.33, RE 4.45) compared to the fairness scores of the two unequal regimes (EU 5.4, RU 4.05).

A quick aside on the two equal regimes. EE (where subjects start the public goods game with a score that is the average of all scores from the skill-based task in a given network) seems to be the less fair of the two equal regimes (for reference, RE assigns an equal score to all subjects that is independent of the skill-based task). This is perhaps not surprising given that EE resembles redistribution in the sense that subjects with higher scores in the skill-based task help raise the scores of the subjects with lower scores, and most Americans are not particularly favorable towards redistribution (McCall 2013).

It is also important to note that while fairness scores are overall higher in the second game (4.62) compared to the first game (4.51), the EU regime actually has a lower average fairness score in the second game (5.53 vs. 5.28). One possible explanation for this drop only in the case of EU is the classic regression to the mean argument: if the fairness score of EU is by chance too high in the first game, it is more likely that it will go down when measurement happens twice. (As can be seen in Table A5.1 below, the coefficient estimate on “Earned x Second” is negative and significant.) A more theoretical, and perhaps more plausible, explanation related to trust is discussed further below; the gist of that argument is that lower levels of cooperation under EU lead to lower levels of trust by the end of the first game, which in turn affect fairness perceptions of the second game.

Figure A5.1 shows the results of fairness perceptions that subjects reported right after being explained which endowment regime they were going to play each public goods game under. At the end of the experiment, subjects were also asked to indicate which of the two endowment regimes they were exposed to was more fair, and which they preferred. The fairness ordering observed in Figure A5.1 is also mostly consistent with these pairwise fairness and preference comparisons that players made at the end of the experiment. EU is chosen as more fair (more preferred) compared to RU by 69% (62%) of players; RE is chosen as more fair (more preferred) compared to RU by 74% (72%) of players; EE is chosen as more fair (more preferred) compared to RU by 64% (60%) of players; EU is chosen as more fair (more preferred) compared to EE by 62% (53%) of players; RE is chosen as more fair (more preferred) compared to EE by 58% (57%) of players; and EU is chosen as more fair (more preferred) compared to RE by 49% (51%) of players. Perhaps the only surprising result here is that while EU has a much higher average fairness score compared to RE (5.4 vs. 4.45), the two regimes are practically considered to be equally fair (preferable) in the pairwise comparisons.

Figure A5.1. Average fairness scores by endowment regime.



Going beyond visual inspection, we can fit a series of models to our data. The general structure of the model with clustered standard errors looks as follows (indices i and j stand for network and subject, respectively), with standard errors clustered at the network (i) level.

$$fairness_{ij} = \beta_0 + \beta_1 earned_i + \beta_2 equal_i + \beta_3 earned_i \times equal_i + \varepsilon_{ij} \quad \text{(A5.1)}$$

An alternative approach to modeling this outcome would be to use a linear mixed-effects model, which explicitly decomposes the error term into two parts.

$$fairness_{ij} = \beta_0 + \beta_1 earned_i + \beta_2 equal_i + \beta_3 earned_i \times equal_i + \xi_i + \varepsilon_{ij} \quad \text{(A5.2)}$$

In both of these models, $fairness_{ij}$ takes discrete values between 1 and 7 (higher values more fair), while $earned_i$ and $equal_i$ each take the values 0 or 1. Note that since subjects play two public goods games in a given session, it is actually possible to fit models to (i) only the first public goods game, (ii) only the second public goods game, or (iii) both public goods games pooled together. While the models described above assume that the model is fit to a single public goods game only (e.g., the first public goods game), it is possible to extend them into the pooled version by adding an additional index k to denote the public goods game (first or second) and a dummy variable denoting whether observations come from the second as opposed to the first game ($second_{ik}$). Such an extension allows us to fit our models to a much larger sample. In this case, standard errors would be clustered at the network (i) and subject (j) levels, while the linear mixed-effects model would include three, instead of two, error terms.

$$fairness_{ijk} = \beta_0 + \beta_1 earned_{ik} + \beta_2 equal_{ik} + \beta_3 earned_{ik} \times equal_{ik} + \beta_4 second_{ik} + \varepsilon_{ijk} \quad \text{(A5.1')}$$

$$fairness_{ijk} = \beta_0 + \beta_1 earned_{ik} + \beta_2 equal_{ik} + \beta_3 earned_{ik} \times equal_{ik} + \beta_4 second_{ik} + \xi_i + \xi_{ij} + \varepsilon_{ijk} \quad \text{(A5.2')}$$

Table A5.1 below presents results from these between-subjects models with fairness as the outcome variable. In addition to the pooled Models A5.1' and A5.2' described above, this table also presents results from a more flexible pooled model that allows for all possible two- and three-way interactions (this model was left out above for the sake of space).

Together, these estimates tell us that being in an earned regime has a large positive direct effect on fairness, whereas being in an equal regime has a still positive but much smaller direct effect. There is also a large negative interaction effect between Earned and Equal, which means that the positive direct effects of Earned and Equal on fairness diminish greatly for the condition EE. The facts that (i) the Earned coefficient is almost four times as large as the Equal coefficient and (ii) the Earned x Equal coefficient is negative suggest that equal arrangements are not necessarily more fair compared to unequal ones. Rather, what makes a regime fair or unfair is the specific mechanism through which equality or inequality comes about, operationalized through the Earned/Random axis. Finally, as can be seen in the flexible pooled model, while the second game is generally higher in terms of fairness (Second is positive), the earned direct effect is actually smaller in the second game (Earned x Second is negative). The models with clustered standard errors and the linear mixed-effects models agree with each other, though the estimates are not identical.

Table A5.1. Fairness perceptions as a function of endowment regimes (between-subjects).

Models with clustered standard errors

	1st game only	2nd game only	Pooled	Pooled (flexible)
Intercept	3.900*** (0.100)	4.208*** (0.095)	4.015*** (0.076)	3.900*** (0.100)
Earned	1.630*** (0.125)	1.072*** (0.135)	1.351*** (0.095)	1.630*** (0.125)
Equal	0.461** (0.153)	0.331* (0.150)	0.402*** (0.105)	0.461** (0.153)
Earned x Equal	-1.673*** (0.192)	-1.267*** (0.199)	-1.468*** (0.143)	-1.673*** (0.192)
Second			0.063 (0.060)	0.308* (0.135)
Earned x Second				-0.558** (0.179)
Equal x Second				-0.130 (0.219)
Earned x Equal x Second				0.406 (0.269)

Linear mixed-effects models

	1st game only	2nd game only	Pooled	Pooled (flexible)
Intercept	3.902*** (0.097)	4.213*** (0.102)	4.022*** (0.073)	3.918*** (0.089)
Earned	1.625*** (0.138)	1.067*** (0.140)	1.360*** (0.090)	1.601*** (0.126)

Equal	0.462** (0.136)	0.328* (0.142)	0.427*** (0.089)	0.481*** (0.123)
Earned x Equal	-1.670*** (0.194)	-1.264*** (0.199)	-1.554*** (0.126)	-1.724*** (0.176)
Second			0.065 (0.054)	0.283* (0.123)
Earned x Second				-0.482** (0.177)
Equal x Second				-0.120 (0.178)
Earned x Equal x Second				0.344 (0.251)

The number of samples (*n*) for the models fit on the first game only, the second game only, and both games pooled together are 1870, 1803, and 3673, respectively, clustered inside 160 sessions. The numbers inside the parentheses are standard errors. Stars denote p-values: ˙ p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

Lastly, it is important to mention two additional points. The first point is related to the possible effect that the difference between a subject's score (in the skill-based task) and their endowment (at the start of a public goods game) might have on that subject's fairness perceptions. Note that a player's endowment may be higher, lower, or identical compared to their score in the skill-based task, depending on the endowment regime. Based on past empirical evidence that shows that winners tend to rationalize their success in moral terms (Ohtsuka and Ohtsuka 2010), there is reason to believe that players who find themselves in a more advantaged position are also more likely to consider their condition to be fair, even if that advantage comes about randomly.

In line with this argument, we see that an increase in score from the skill-based task to the public goods game leads to an increase in fairness perceptions (an average increase of ≈0.2 in fairness per every 100 points, significant at the $p<0.001$ level). The inclusion of this additional predictor does not change the other estimates much, and the previous conclusions stand. Results also remain unchanged if an additional predictor for experience is included in the models. (Between-subjects and within-subjects results agree on all of these.)

The second point is related to network-level analyses. While all of the analyses above were conducted on a dataset where there is a row for every network-subject(-game), it is also possible to conduct a more crude analysis using an aggregate version of the data where rows are per network(-game).

$$avg\_fairness_i = \beta_0 + \beta_1 earned_i + \beta_2 equal_i + \beta_3 earned_i \times equal_i + \varepsilon_i \quad \text{(A5.1'')}$$

$$avg\_fairness_{ik} = \beta_0 + \beta_1 earned_{ik} + \beta_2 equal_{ik} + \beta_3 earned_{ik} \times equal_{ik} + \beta_4 second_{ik} + \varepsilon_{ik} \quad \text{(A5.1''')}$$

$$avg\_fairness_{ik} = \beta_0 + \beta_1 earned_{ik} + \beta_2 equal_{ik} + \beta_3 earned_{ik} \times equal_{ik} + \beta_4 second_{ik} + \xi_i + \varepsilon_{ik} \quad \text{(A5.2''')}$$

Results from such an aggregate analysis do not change the conclusions above (this is also the case for aggregate within-subjects models). Therefore, separate tables for these models are not included here.

Cooperation patterns

Figure A5.2 below visualizes overall patterns of cooperation across rounds. As can be seen in this figure, (i) cooperation decreases in later rounds, and (ii) cooperation levels are lower in the second game compared to the first one. The first pattern is one that is widely observed in similar networked games (Mason, Suri, and Watts 2014); in fact, given that the ratio of benefit of cooperation (*b=100*) to cost of cooperation (*c=50*) is less than the average number of connections (*k=6.8*) in this case (100/50=2 < 6.8), this pattern is both theoretically expected and empirically shown in Rand et al 2014. (While Figure A5.2 presents aggregate results, a look into the cooperation histories of individual subjects across rounds are mostly in line with the pattern of gradual collapse of cooperation over time. In fact, around 60% of all cooperation histories are strictly non-increasing, that is, once a subject starts defecting, they never cooperate again.)

The second pattern can be explained in reference to the fact that by the end of the first game, subjects most likely already lost some amount of trust in other subjects after observing at least some of them defect. In fact, if we regress average cooperation in the second game on average trust at the end of the first game, controlling for average cooperation in the first game, the coefficient estimate on average trust is 0.073*** (0.017), which can be interpreted to mean that a one unit increase in average trust (on a scale of 1 to 7) at the end of the first game leads to a 7% increase in average cooperation in the second game. Furthermore, given that average trust at the end of the first game is very strongly correlated with average cooperation in the first game ($\varrho = 0.91$), the argument that lower levels of cooperation in the second game is partially due to the generally negative impact of the first game on trust becomes more plausible.

Figure A5.2. Overall patterns of cooperation.

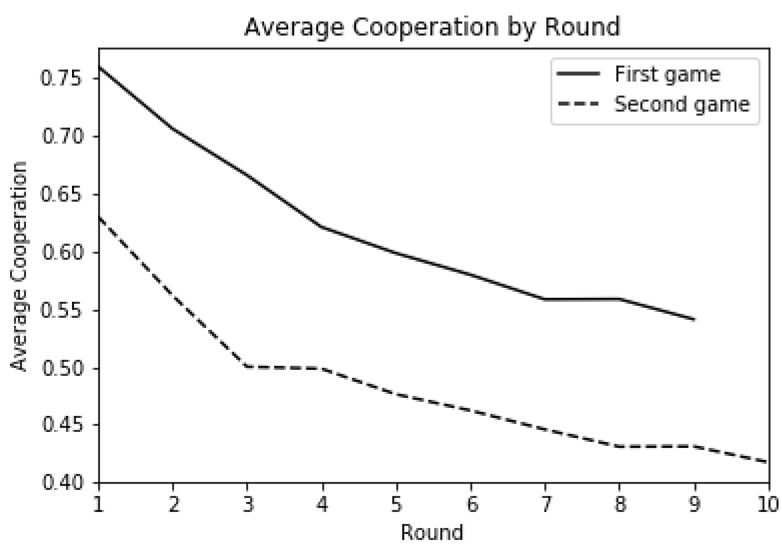
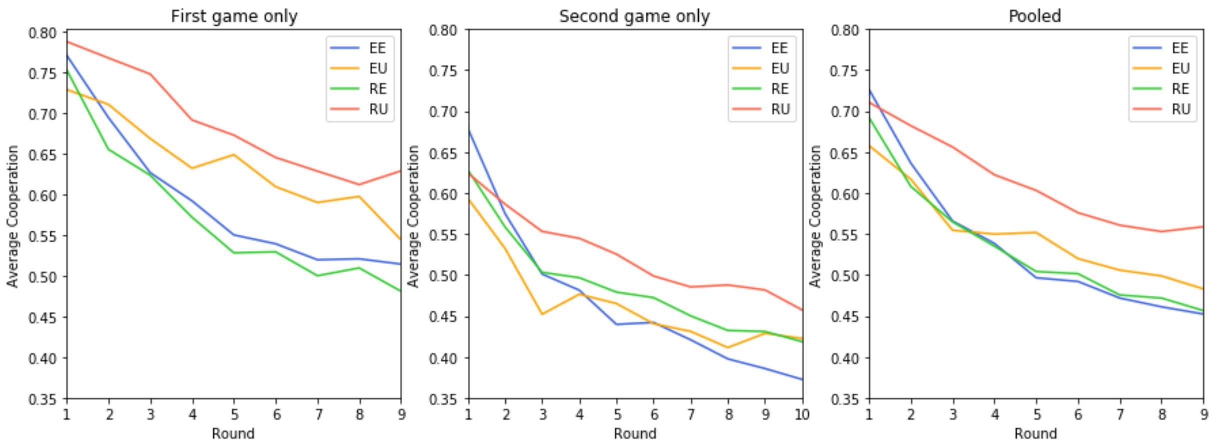


Figure A5.3 below visualizes patterns of cooperation across rounds in different endowment regimes. As can be seen in this figure, average cooperation in EU is consistently lower (~5%) compared to RU, while patterns of cooperation in EE and RE are a lot closer to one another (EE is slightly higher compared to RE in the first game, while RE is generally higher compared to EE in the second game). Another observation to make is that the two unequal regimes (EU and RU) are overall higher compared to the two equal regimes (EE and RE), especially in the first game. These patterns suggest not only a negative "equal" effect, giving support to the argument that people prefer unequal societies (Norton 2014), but also a negative "earned" effect, whereby players are less willing to share their wealth with others if they believe to have "earned" that wealth, which would be consistent with an entitlement effect (Krawczyk 2010). (For reference, average levels of cooperation across regimes in the first game are: RE 57%, EE 59%, EU 64%, and RU 69%; and average levels of cooperation across regimes in the second game are: EU 47%, EE 47%, RE 49%, and RU 52%.)

Figure A5.3. Average cooperation by round by endowment regime.



Similar to the approach we took above when discussing fairness perceptions, we can go beyond visual inspection and fit a series of models to our data with cooperation as the outcome. The indices i, j, k, and l are for session, player, game, and round, respectively.

The between-subjects models with standard errors clustered at the network (i) and subject (j) levels can be written down as:

$$g(cooperation_{ijl}) = \beta_0 + \beta_1 earned_i + \beta_2 equal_i + \beta_3 earned_i \times equal_i + \sum_{l=1}^{R} \beta_{4l}\, round_{il} + \varepsilon_{ijl} \quad \text{(A5.3)}$$

$$g(cooperation_{ijkl}) = \beta_0 + \beta_1 earned_{ik} + \beta_2 equal_{ik} + \beta_3 earned_{ik} \times equal_{ik} + \sum_{l=1}^{R} \beta_{4l}\, round_{ikl} + \beta_5 second_{ik} + \varepsilon_{ijkl} \quad \text{(A5.3')}$$

Similarly, the (generalized) linear mixed-effects between-subjects models are:

$$g(cooperation_{ijl}) = \beta_0 + \beta_1 earned_i + \beta_2 equal_i + \beta_3 earned_i \times equal_i + \sum_{l=1}^{R} \beta_{4l}\, round_{il} + \xi_i + \xi_{ij} + \varepsilon_{ijl} \quad \text{(A5.4)}$$

$$g(cooperation_{ijkl}) = \beta_0 + \beta_1 earned_{ik} + \beta_2 equal_{ik} + \beta_3 earned_{ik} \times equal_{ik} + \sum_{l=1}^{R} \beta_{4l}\, round_{ikl} + \beta_5 second_{ik} + \xi_i + \xi_{ij} + \xi_{ijk} + \varepsilon_{ijkl} \quad \text{(A5.4')}$$

$g(.)$ is the link function used to model the outcome, which is logistic in the case of both $cooperation_{ijl}$ and $cooperation_{ijkl}$. (Conclusions do not change if ordinary least squares is used instead.)

Table A5.2 below presents results from these between-subjects models with cooperation decision as the outcome variable. As can be seen in this table, there is a fair amount of variation between the estimates returned by the models with clustered standard errors on the one hand, and the generalized linear mixed-effects models on the other, especially in the case of pooled estimates. To begin with, while results from the "1st game only" and "2nd game only" models are similar in terms of sign and significance across both models, estimates from the mixed-effects models are usually more than twice as large in magnitude. Regardless, both models show a larger and significant negative Equal effect and a smaller and insignificant negative Earned effect in the first game, while the Earned effect is larger than Equal in the second game, though neither of them is significant. In other words, between-subjects models indicate that there is less cooperation under equality in the first game, while cooperation across regimes are not significantly different in the second game.

The divergence between the two sets of models becomes more stark in the case of the pooled estimates. While the models with clustered standard errors continue showing a larger and significant negative Equal effect (and a smaller and insignificant negative Earned effect) in the pooled models, estimates from the linear mixed-effects models indicate a larger and significant Earned effect (and a smaller and possibly insignificant negative Equal effect), while also showing a significant positive Equal x Earned interaction. In other words, the mixed-effects models flip the story and attribute the larger effect to Earned, while Equal could still have an effect, though smaller. The results are robust to the inclusion of additional predictors for player score and experience in the models.

Table A5.2. Cooperation decision as a function of endowment regimes (between-subjects).

Models with clustered standard errors

	1st game only	2nd game only	Pooled	Pooled (flexible)
Intercept	1.453*** (0.146)	0.692*** (0.155)	1.331 (0.114)	1.401*** (0.142)
Earned	-0.230 (0.210)	-0.240 (0.225)	-0.247 (0.154)	-0.229 (0.209)
Equal	-0.503* (0.194)	-0.152 (0.229)	-0.327* (0.154)	-0.501* (0.193)
Earned x Equal	0.312 (0.284)	0.170 (0.337)	0.259 (0.209)	0.310 (0.283)
Second			-0.530*** (0.038)	-0.668*** (0.188)
Earned x Second				-0.024 (0.311)
Equal x Second				0.349 (0.296)
Earned x Equal x Second				-0.116 (0.464)

Generalized linear mixed-effects models

	1st game only	2nd game only	Pooled	Pooled (flexible)
Intercept	3.447*** (0.293)	1.841*** (0.333)	2.820*** (0.157)	2.799*** (0.164)
Earned	-0.519 (0.394)	-0.587 (0.447)	-0.426*** (0.071)	-0.443*** (0.102)
Equal	-1.136** (0.391)	-0.485 (0.451)	-0.182** (0.070)	-0.138 (0.098)

Earned x Equal	0.734 (0.556)	0.673 (0.636)	0.331** (0.098)	0.357* (0.139)
Second			-1.166*** (0.036)	-1.124*** (0.095)
Earned x Second				0.033 (0.139)
Equal x Second				-0.092 (0.140)
Earned x Equal x Second				-0.043 (0.197)

The number of samples (n) for the models fit on the first game only, the second game only, and both games pooled together are 16579, 17867, and 32677, clustered inside 160 sessions and 1870, 1803, and 1870 players, respectively. All models control for round by including round dummies as predictors. Estimates are left in the original log-odds scale. The numbers inside the parentheses are standard errors. Stars denote p-values: ˙ p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

While the above models were fit on an subject-level dataset with the structure network-subject(-game)-round per row, once again, we can fit our models on a more aggregate dataset with the structure network(-game)-round, where the outcome would now be average cooperation in a given round (continuous), rather than the cooperation decision for a specific subject (0, 1). As discussed in the next paragraph, the main advantage of the network-level models is easier interpretability.

$$avg\_coop_{il} = \beta_0 + \beta_1 earned_i + \beta_2 equal_i + \beta_3 earned_i \times equal_i + \sum_{l=1}^{R} \beta_{4l}\, round_{il} + \varepsilon_{il} \quad \text{(A5.3'')}$$

$$avg\_coop_{ikl} = \beta_0 + \beta_1 earned_{ik} + \beta_2 equal_{ik} + \beta_3 earned_{ik} \times equal_{ik} + \sum_{l=1}^{R} \beta_{4l}\, round_{ikl} + \beta_5 second_{ik} + \varepsilon_{ikl} \quad \text{(A5.3''')}$$

$$avg\_coop_{il} = \beta_0 + \beta_1 earned_i + \beta_2 equal_i + \beta_3 earned_i \times equal_i + \sum_{l=1}^{R} \beta_{4l}\, round_{il} + \xi_i + \varepsilon_{il} \quad \text{(A5.4'')}$$

$$avg\_coop_{ikl} = \beta_0 + \beta_1 earned_{ik} + \beta_2 equal_{ik} + \beta_3 earned_{ik} \times equal_{ik} + \sum_{l=1}^{R} \beta_{4l}\, round_{ikl} + \beta_5 second_{ik} + \xi_i + \xi_{ik} + \varepsilon_{ikl} \quad \text{(A5.4''')}$$

Table A5.3 below presents results from these network-level models. The estimates mostly mirror those in Table A5.2 but are much easier to interpret given that the outcome (average cooperation) is now continuous: the models with clustered standard errors (and the linear mixed-effects models for each game separately) indicate a significant large negative Equal effect corresponding up to a 10% decrease in cooperation, while the pooled linear mixed-effects models show a moderately sized (~4%) significant negative Earned effect. (Note that this 4% estimate is very close to the within-subjects Earned estimate.)

Taken together, these results do not allow us to reach a definitive answer as to whether the main between-subjects effect is due to Equal or Earned, though there is evidence to believe that both of these axes likely have a non-negligible effect on players' behavior: in particular, if we go back to Figure A5.3, we can see that while the Equal axis seems to have a clear negative effect on cooperation in the first game, the persistent difference between EU and RU in both games seems to be the main driver of the estimated Earned effect.

Table A5.3. Average cooperation as a function of endowment regimes (between-subjects).

Models with clustered standard errors

	1st game only	2nd game only	Pooled	Pooled (flexible)
Intercept	0.818*** (0.029)	0.676*** (0.037)	0.808*** (0.025)	0.816*** (0.030)
Earned	-0.051 (0.047)	-0.069 (0.054)	-0.061$^{\cdot}$ (0.036)	-0.051 (0.047)
Equal	-0.103* (0.045)	-0.059 (0.058)	-0.081* (0.037)	-0.103* (0.045)
Earned x Equal	0.064 (0.066)	0.074 (0.083)	0.072 (0.052)	0.064 (0.066)
Second			-0.121*** (0.009)	-0.136** (0.042)
Earned x Second				-0.021 (0.072)
Equal x Second				0.044 (0.071)
Earned x Equal x Second				0.017 (0.108)

Linear mixed-effects models

	1st game only	2nd game only	Pooled	Pooled (flexible)
Intercept	0.818*** (0.034)	0.676*** (0.042)	0.773*** (0.020)	0.767*** (0.021)
Earned	-0.051 (0.046)	-0.069 (0.058)	-0.039*** (0.009)	-0.042** (0.013)
Equal	-0.103* (0.046)	-0.059 (0.058)	-0.005 (0.009)	0.004 (0.013)
Earned x Equal	0.064 (0.065)	0.074 (0.083)	0.015 (0.013)	0.025 (0.018)
Second			-0.121*** (0.004)	-0.111*** (0.012)
Earned x Second				0.007 (0.018)
Equal x Second				-0.018 (0.018)
Earned x Equal x Second				-0.021 (0.025)

The number of samples (*n*) for the models fit on the first game only, the second game only, and both games pooled together are 1440, 1600, and 2880, respectively, clustered inside 160 sessions. All models control for round by including round dummies as predictors. The numbers inside the parentheses are standard errors. Stars denote p-values: $^{\cdot}$ p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

Finally, it is important to note that while other studies found a null direct effect of level of inequality in a similar networked setting (Nishi et al 2015), none of the endowment regimes here (not even random equal) are directly comparable to that setting given that in this case score allocation is always preceded by a skill-based task, which ensures that subjects in all regimes must first engage in an activity that requires an effort on their part. (Earlier studies conducted in much smaller groups of three to four subjects remain largely inconclusive: Chan et al (1999) found a negative equality effect, Cherry, Kroll, and Shogren (2005) found a positive equality effect, and Sadrieh and Verbon (2006) found a null equality effect. For reference, Cherry, Kroll, and Shogren (2005) also found that source of endowment does not make a difference, though our experimental setup is too different from theirs to allow for a direct comparison.)

Citations

Chan, K., S. Mestelman, R. Moir, and R. Muller. 1999. “Heterogeneity and the Voluntary Provision of Public Goods.” *Experimental Economics* 2: 5-30.

Cherry, T., S. Kroll, and J. Shogren. 2005. “The impact of endowment heterogeneity and origin on public good contributions: evidence from the lab.” *Journal of Economic Behavior & Organization* 57(3): 357-365.

Krawczyk, M. 2010. “A glimpse through the veil of ignorance: Equality of opportunity and support for redistribution.” *Journal of Public Economics* 94: 131-141.

Mason, W., S. Suri, and D. Watts. 2014. “Long-Run Learning in Games of Cooperation.” *EC’14*, June 8-12, Stanford, CA, USA.

McCall, L. 2013. *The Undeserving Rich American Beliefs about Inequality, Opportunity, and Redistribution.* New York: Cambridge University Press.

Nishi, A., H. Shirado, D. Rand, and N. Christakis. 2015. “Inequality and visibility of wealth in experimental social networks.” *Nature* 526: 426-429.

Norton, M. 2014. “Unequality: Who Gets What and Why It Matters.” *Policy Insights from the Behavioral and Brain Sciences* 1(1): 151-155.

Ohtsuka, K. and T. Ohtsuka. 2010. “Vietnamese Australian Gamblers’ Views on Luck and Winning: Universal Versus Culture-specific Schemas.” *Asian Journal of Gambling Issues and Public Health* 1(1): 34-46.

Rand, D., M. Nowak, J. Fowler, and N. Christakis. 2014. “Static network structure can stabilize human cooperation.” *PNAS* 111(48): 17093-17098.

Sadrieh, A. and H. Verbon 2006. “Inequality, cooperation, and growth: An experimental study.” *European Economic Review* 50(5): 1197-1222.

Appendix 6: Tie formation/breakage models

Figures A6.1 and A6.2 visualize tie formation patterns across games, rounds, and endowment regimes. Table A6.1 presents results from the subject-level between-subjects tie formation models. The upper panel presents estimates not controlling for alter's cooperation choice, while the lower panel presents estimates controlling for alter's cooperation choice. Both sets of models use clustered standard errors. The network-level counterparts of these models are presented in Table A6.2. Figure A6.3 visualizes the effect of changing from one endowment regime to another on tie formation. Table A6.3 presents results from the within-subjects tie formation models. The upper panel presents estimates from the subject-level models, while the lower panel presents estimates from the network-level models. Both sets of models use clustered standard errors. Figures A6.4, A6.5, and A6.6 and Tables A6.4, A6.5, and A6.6 repeat the above analyses with outcome as tie breakage.

Table A6.1. Add tie choice as a function of endowment regimes (between-subjects).

Models not controlling for alter's cooperation choice

	1st game only	2nd game only	Pooled	Pooled (flexible)
Intercept	1.191 (0.116)***	1.012 (0.114)***	1.183 (0.081)***	1.187 (0.098)***
Earned	-0.089 (0.120)	-0.258 (0.109)*	-0.189 (0.075)*	-0.088 (0.120)
Equal	-0.198 (0.102)˙	-0.086 (0.112)	-0.136 (0.071)˙	-0.197 (0.101)
Earned x Equal	0.188 (0.157)	0.330 (0.148)*	0.276 (0.096)**	0.187 (0.157)
Second			-0.167 (0.041)***	-0.181 (0.124)
Earned x Second				-0.165 (0.188)
Equal x Second				0.127 (0.176)
Earned x Equal x Second				0.139 (0.253)

Models controlling for alter's cooperation choice

	1st game only	2nd game only	Pooled	Pooled (flexible)
Intercept	-0.146 (0.121)	-0.161 (0.143)	-0.223 (0.097)*	-0.250 (0.102)*
Earned	0.029 (0.099)	-0.181 (0.114)	-0.078 (0.073)	0.041 (0.100)
Equal	0.008 (0.090)	-0.017 (0.111)	0.015 (0.071)	0.029 (0.092)
Earned x Equal	0.049 (0.130)	0.331 (0.152)*	0.187 (0.102)˙	0.036 (0.132)
Alter's coop choice	1.642 (0.082)***	1.932 (0.090)***	1.795 (0.075)***	1.794 (0.075)***
Second			0.075 (0.041)˙	0.131 (0.106)
Earned x Second				-0.207 (0.157)

Equal x Second				-0.030 (0.151)
Earned x Equal x Second				0.272 (0.203)

The number of samples (*n*) for the models fit on the first game only, the second game only, and both games pooled together are 12826, 15205, and 26100, clustered inside 160 sessions and 1832, 1785, and 1858 players, respectively. All models control for round by including round dummies as predictors. Estimates are left in the original log-odds scale. The numbers inside the parentheses are standard errors. Stars denote p-values: ˙ p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

Table A6.2. Average add choice as a function of endowment regimes (between-subjects).

Models not controlling for alter's cooperation choice

	1st game only	2nd game only	Pooled	Pooled (flexible)
Intercept	0.774 (0.023)	0.742 (0.025)***	0.776 (0.018)***	0.767 (0.021)***
Earned	0.003 (0.029)	-0.076 (0.027)**	-0.038 (0.019)*	0.003 (0.029)
Equal	-0.040 (0.025)	-0.040 (0.028)	-0.039 (0.018)*	-0.040 (0.025)
Earned x Equal	0.028 (0.037)	0.112 (0.038)**	0.071 (0.025)**	0.028 (0.037)
Second			-0.036 (0.010)***	-0.019 (0.027)
Earned x Second				-0.080 (0.044)˙
Equal x Second				0.003 (0.041)
Earned x Equal x Second				0.085 (0.059)

Models controlling for alter's cooperation choice

	1st game only	2nd game only	Pooled	Pooled (flexible)
Intercept	0.457 (0.031)***	0.481 (0.030)***	0.457 (0.025)***	0.450 (0.026)***
Earned	0.018 (0.023)	-0.048 (0.021)*	-0.015 (0.015)	0.018 (0.023)
Equal	0.002 (0.021)	-0.006 (0.022)	0.000 (0.015)	0.002 (0.021)
Earned x Equal	0.003 (0.029)	0.075 (0.030)*	0.038 (0.021)˙	0.003 (0.029)
Alter's coop choice	0.376 (0.028)***	0.370 (0.026)***	0.381 (0.022)***	0.379 (0.022)***
Second			0.012 (0.009)	0.030 (0.022)
Earned x Second				-0.065 (0.033)*
Equal x Second				-0.006 (0.032)
Earned x Equal x Second				0.070 (0.043)

The number of samples (*n*) for the models fit on the first game only, the second game only, and both games pooled together are 1231, 1366, and 2449, respectively, clustered inside 160 sessions. All models control for round. The numbers inside the parentheses are standard errors. Stars denote p-values: ˙ p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

Table A6.3. Tie formation as a function of endowment regimes (within-subjects).

Subject-level models

	Not controlling for alter's coop choice	Controlling for alter's coop choice
Intercept	-0.047 (0.033)	0.013 (0.029)
Δ Earned	-0.013 (0.023)	-0.006 (0.021)
Δ Equal	0.027 (0.021)	0.019 (0.018)
Δ Earned x Equal	0.017 (0.030)	0.029 (0.029)
Δ Alter's coop choice		0.377 (0.021)***

Network-level models

	Not controlling for alter's coop choice	Controlling for alter's coop choice
Intercept	-0.050 (0.025)*	0.005 (0.023)
Δ Earned	-0.022 (0.016)	-0.012 (0.015)
Δ Equal	-0.019 (0.017)	-0.012 (0.015)
Δ Earned x Equal	0.035 (0.023)	0.027 (0.023)
Δ Alter's coop choice		0.397 (0.034)***

The number of samples (*n*) for the individual-level models fit is 13171, clustered inside 1558 players and 160 sessions. The number of samples (*n*) for the session-level models fit is 1184, clustered inside 160 sessions. The numbers inside the parentheses are standard errors. Stars denote p-values: ˙ p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

Table A6.4. Cut tie choice as a function of endowment regimes (between-subjects).

Models not controlling for alter's cooperation choice

	1st game only	2nd game only	Pooled	Pooled (flexible)
Intercept	-1.093 (0.151)***	-0.763 (0.123)***	-1.064 (0.107)***	1.099 (0.143)***
Earned	-0.088 (0.169)	0.177 (0.146)	0.060 (0.102)	-0.087 (0.169)
Equal	0.385 (0.151)*	0.057 (0.153)	0.235 (0.113)*	0.385 (0.150)*
Earned x Equal	-0.079 (0.220)	-0.226 (0.224)	-0.165 (0.135)	-0.079 (0.220)
Second			0.268 (0.042)***	0.343 (0.152)*
Earned x Second				0.237 (0.248)

Equal x Second				-0.323 (0.208)
Earned x Equal x Second				-0.102 (0.355)

Models controlling for alter's cooperation choice

	1st game only	2nd game only	Pooled	Pooled (flexible)
Intercept	0.430 (0.137)**	0.491 (0.119)***	0.476 (0.099)***	0.505 (0.118)***
Earned	-0.248 (0.113)*	0.108 (0.115)	-0.067 (0.079)	-0.261 (0.111)*
Equal	0.132 (0.110)	0.087 (0.129)	0.106 (0.084)	0.118 (0.109)
Earned x Equal	0.159 (0.152)	-0.293 (0.163)$^{\cdot}$	-0.042 (0.111)	0.175 (0.150)
Alter's coop choice	-1.993 (0.091)***	-2.192 (0.096)***	-2.108 (0.083)***	-2.106 (0.083)***
Second			0.008 (0.041)	-0.057 (0.112)
Earned x Second				0.347 (0.165)*
Equal x Second				-0.026 (0.162)
Earned x Equal x Second				-0.393 (0.225)$^{\cdot}$

The number of samples (n) for the models fit on the first game only, the second game only, and both games pooled together are 12348, 12365, and 23482, clustered inside 160 sessions and 1842, 1784, and 1857 players, respectively. All models control for round by including round dummies as predictors. Estimates are left in the original log-odds scale. The numbers inside the parentheses are standard errors. Stars denote p-values: $^{\cdot}$ p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

Table A6.5. Average cut choice as a function of endowment regimes (between-subjects).

Models not controlling for alter's cooperation choice

	1st game only	2nd game only	Pooled	Pooled (flexible)
Intercept	0.256 (0.028)***	0.285 (0.026)***	0.239 (0.021)***	0.249 (0.026)***
Earned	-0.021 (0.030)	0.070 (0.033)*	0.026 (0.022)	-0.021 (0.030)
Equal	0.066 (0.029)*	0.055 (0.037)	0.061 (0.025)*	0.066 (0.029)*
Earned x Equal	-0.004 (0.043)	-0.099 (0.051)$^{\cdot}$	-0.050 (0.032)	-0.004 (0.043)
Second			0.060 (0.010)***	0.041 (0.028)
Earned x Second				0.094 (0.049)$^{\cdot}$
Equal x Second				-0.010 (0.045)
Earned x Equal x Second				-0.091 (0.072)

Models controlling for alter's cooperation choice

	1^{st} game only	2^{nd} game only	Pooled	Pooled (flexible)
Intercept	0.644 (0.029)***	0.628 (0.026)***	0.641 (0.023)***	0.652 (0.024)***
Earned	-0.047 (0.018)*	0.026 (0.022)	-0.010 (0.015)	-0.048 (0.018)**
Equal	0.010 (0.020)	0.016 (0.024)	0.014 (0.016)	0.009 (0.019)
Earned x Equal	0.039 (0.028)	-0.049 (0.032)	-0.001 (0.024)	0.040 (0.028)
Alter's coop choice	-0.482 (0.024)***	-0.491 (0.023)***	-0.494 (0.019)***	-0.493 (0.019)***
Second			-0.002 (0.009)	-0.025 (0.019)
Earned x Second				0.077 (0.029)**
Equal x Second				0.010 (0.028)
Earned x Equal x Second				-0.081 (0.038)*

The number of samples (*n*) for the models fit on the first game only, the second game only, and both games pooled together are 1279, 1434, and 2554, respectively, clustered inside 160 sessions. All models control for round. The numbers inside the parentheses are standard errors. Stars denote p-values: ˙ p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

Table A6.6. Tie breakage as a function of endowment regimes (within-subjects).

Subject-level models

	Not controlling for alter's coop choice	Controlling for alter's coop choice
Intercept	0.053 (0.031)˙	0.008 (0.028)
Δ Earned	0.006 (0.021)	-0.019 (0.019)
Δ Equal	-0.044 (0.024)˙	-0.050 (0.020)*
Δ Earned x Equal	0.022 (0.035)	0.028 (0.031)
Δ Alter's coop choice		-0.433 (0.024)***

Network-level models

	Not controlling for alter's coop choice	Controlling for alter's coop choice
Intercept	0.043 (0.024)˙	-0.005 (0.021)
Δ Earned	0.033 (0.020)˙	-0.0003 (0.016)
Δ Equal	0.014 (0.021)	0.0007 (0.019)
Δ Earned x Equal	-0.020 (0.029)	0.0020 (0.027)
Δ Alter's coop choice		-0.489 (0.031)***

The number of samples (n) for the individual-level models fit is 10083, clustered inside 1540 players and 160 sessions. The number of samples (n) for the session-level models fit is 1274, clustered inside 160 sessions. The numbers inside the parentheses are standard errors. Stars denote p-values: $^{\cdot}$ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

Figure A6.1. Overall patterns of tie formation.

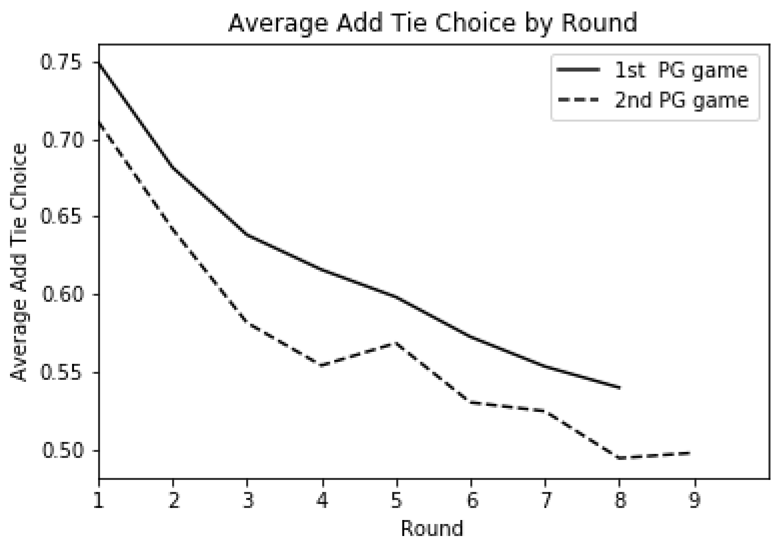


Figure A6.2. Average tie formation by round by endowment regime.

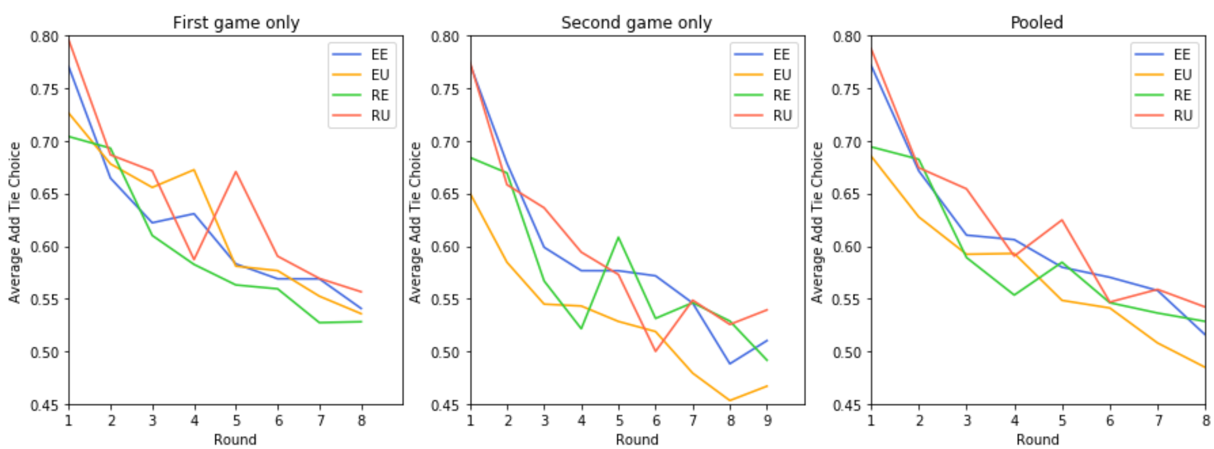


Figure A6.3. Heatmap of change in add tie choice between games.

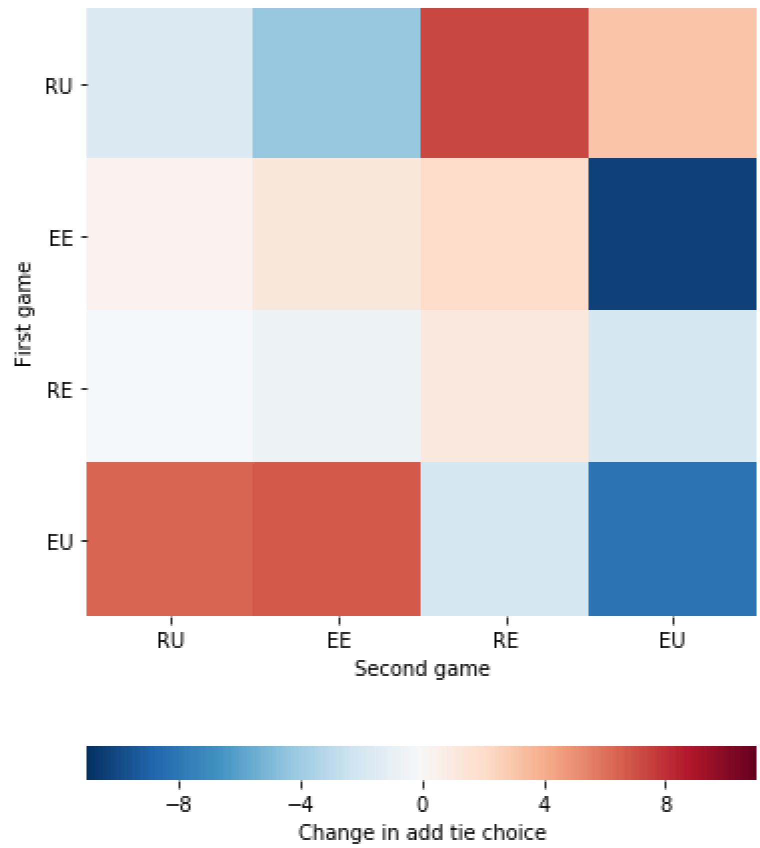


Figure A6.4. Overall patterns of tie breakage.

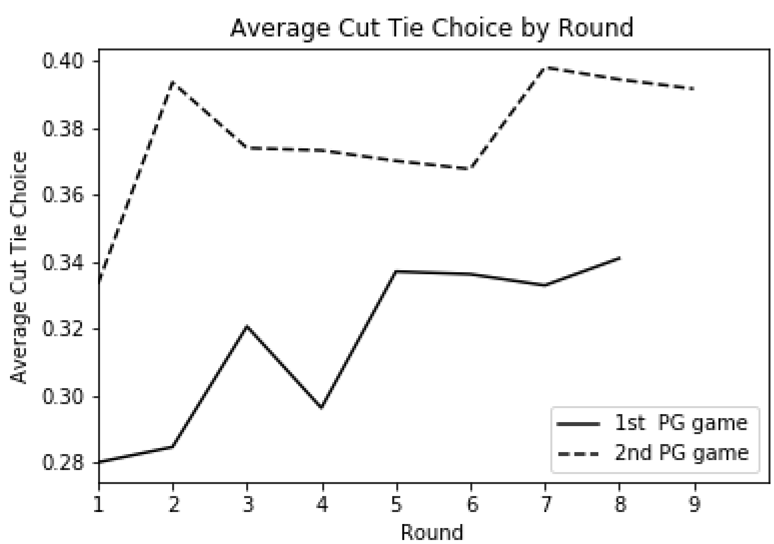


Figure A6.5. Average tie breakage by round by endowment regime.

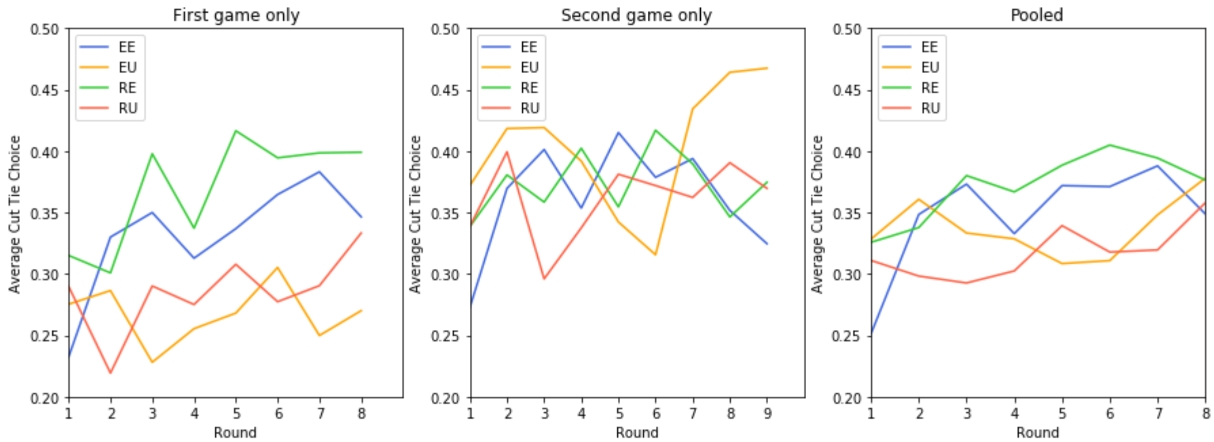


Figure A6.6. Heatmap of change in cut tie choice between games.