Reciprocating preferences in two-sided matching: An experimental investigation

Pre-analysis plan

Timm Opitz[1,2] Christoph Schwaiger[1]

[1]Department of Economics, University of Munich (LMU), Germany

[2]Max Planck Institute for Innovation and Competition, Munich, Germany

1 Introduction

Matching markets describe settings where the efficient formation of (mutually) beneficial relationships does not rely on the coordination function of prices. Examples include *dating markets*, *labor markets* and *university admissions*. Under the deferred acceptance mechanism (DA) (Gale and Shapley (1962)), formed matches ought to be stable. No player prefers to remain single instead of being with her current partner and no pair of players would rather be matched with each other compared to their current partner. While this result builds on the assumptions of strict and invariable preference orders under complete information, empirical evidence for whether this is a good approximation to real-word matching markets is lacking.

Building on the theoretical framework of Opitz and Schwaiger (2021), we test whether agents' preferences are subject to the information about others' preferences. More specifically, we analyze whether agents change their preference order once they know how other players ranked them and document how this affects stability of the DA. We hypothesize that players display a preference for being ranked high on the preference lists. Hence, being preferred as a partner by someone leads to a more favorable evaluation of this individual. This may be both due to a (psychological) value of being liked and different expectations of other's behavior conditional based on whether one was a popular or unpopular option. Thinking about a partner-choice setting, for an individual it may be genuinely important that their partner likes them as well. In labor markets, individuals may prefer to work for a company that is also very eager to work with them because they hope to receive more support and have better career opportunities. We study this question in a laboratory experiment, being able to cleanly manipulate the information sets of players and analyze preference changes associated with this information. We contribute to the growing literature on experimental matching markets (see Hakimov and Kübler (2020) for a review). Closest to the structure of this experiment is previous work on the effect of information about own and others preference profiles in decentralized markets (Haruvy & Ünver, 2007; Pais, Pintér, & Veszteg, 2012), one-sided centralized (Pais & Pintér, 2008) as well as two-sided centralized markets (Pais, Pintér, & Veszteg, 2011). While these papers explicitly analyze whether agents use the additional information to misrepresent their preference orders strategically, we are interested in the causal effect of knowing one's own rank in the preference order of potential partners.

In the experiment, we let participants form teams for a subsequent Public Goods Game (PGG) through a centralized matching mechanism. Participants indicate with whom they would like to play the PGG by submitting a rank-ordered list. After submitting initial rank-ordered lists, a subset of players receives the information how they were ranked (by the other market side the subjects had to rank before) and are allowed to adapt their preference orders. Those potential changes in the preference orders and the associated changes in the matching outcome of the DA constitute our first set of outcome variables. The second set of outcome variables is based on behavior in the PGG. Inter alia, this allows us to shed light on potential mechanisms that drive the preference changes, including belief-based and preference-based explanations.

2 Experimental Design

The experiment consists of a team-formation process and a PGG that is played within the formed dyads. Teams are formed through a centralized matching mechanism. The underlying preferences of players are based on self-reported questionnaire information of the potential partners. After being matched with one of the potential partners, participants play the PGG with the (known) partner.

We compare behavior under two information structures in a between-subject design.[1] In the baseline condition (*No-Info*), participants never know how their potential (and actual) partners rank them. In the treatment condition (*Info*), participants do receive the information how they are ranked before submitting their final preference list. This allows them to incorporate this information into their own preferences and gives them the option to adjust behavior in the PGG based on their knowledge of how much their partner wanted to do be matched with them.

[1] More specifically, we compare within-subject changes across two conditions which makes parts of the design more akin to a traditional diff-in-diff setting with repeated measures.

During the team-formation process, players interact within *matching groups*. We study a setting of two-sided matching in a one-to-one market. Hence, half of the players within each *matching group* take the role of proposers, half the role of receivers. Within each experimental session, there will be multiple matching groups, each consisting of 8 participants. To increase statistical power, we reshuffle matching groups 4 times. No proposer will interact with the same receiver twice and vice versa.

For simplicity, the following description of all experimental design stages reads as if the experiment was a one-shot game. After this, we give a detailed explanation of the repeated nature. The design is also visualized in Figure 1.

Questionnaires [Stage 1]

- Each participant fills out a questionnaire with 15 questions.
- Five questions with a 4 point Likert scale in the following categories:
  - Personality questions (e.g. Big Five)
  - Preferred leisure activities
  - Societal opinions

Matching [Stage 2 & 3]

- Participants are informed about the upcoming PGG (incl. description/instructions).
- Participants are informed that they can can indicate preferences for their partner based on the information in the previously answered questionnaires.
- Participants are explained the basic properties of the DA mechanism.
- Questionnaires are shared between the participants. Proposers get five randomly selected answers from each receiver; receivers get five randomly selected (but mutually exclusive) answers from each proposer.
- The reason for sharing distinct questions is to minimize the initial correlations between preferences. The more the initial preferences are correlated, the less we can draw inferences based on our treatment.
- Based on the questionnaires, participants rank the agents from the other side of the market in terms of the desirability to play the PGG with them.
- After participants have submitted their rank ordered lists, the DA mechanism implements the (provisional) rankings.

Rematching [Stage 4]

[Only proposers enter this stage. Receivers are not informed about this stage.]

- Treatment Variation
  - *No-Info*: Proposers see with whom they have been matched.

  - *Info*: Proposers see with whom they have been matched. In addition, they see how all receivers ranked them.
- Proposers know that receivers will only know with whom they are matched in the end and that receivers will never receive any information on the proposers' preferences (and changes of preferences).
- Proposers submit a (potentially revised) preference list to the DA mechanism.
- One of the two decisions of proposers is randomly implemented to determine the final matching. This randomized procedure guarantees that both the initial submission, as well as the potentially revised preference order are incentive compatible.
- Receivers do not play an active role at this stage of the experiment. Their preferences remain fixed and they do not receive any information.

Public Goods Game [Stage 5]

- "ABC-framework of cooperation" (Gächter, Kölle, & Quercia, 2017), based on Fischbacher, Gächter, and Fehr (2001), but additionally includes an elicitation of beliefs.
- Two-player PGG with a marginal per capita return (MPCR) to 0.75 (comparable to previous literature on two-player PGG, e.g. Goeree, Holt, and Laury (2002); Spraggon and Oxoby (2009)). Hence, free-riding is the dominant strategy from an individual perspective. However, since the sum of marginal returns is larger than 1, contributing the entire endowment maximizes the group surplus.
- **Unconditional and Conditional Contributions** Proposers state both their conditional and unconditional contributions (Fischbacher et al., 2001), knowing that one of the two will be randomly chosen to be payoff relevant. First, they state their unconditional contribution. Second, proposers fill out the conditional contribution table. Receivers only state their unconditional contribution.[2]
- **Belief Elicitation** We ask both proposers and receivers for their point belief about their matched partner's contribution (incentivized).

Additional Measures [Stage 6]

- **Cognitive ability:** Cognitive ability may influence the extent to which proposers treat the preferences of receivers as influential for their own (adjusted) preferences. Proposers with higher cognitive ability may be more likely to believe that receivers who ranked them high in their preference lists contribute more and accordingly adjust their behavior in the rematching

[2] This circumvents the problem with conditional contributions already raised in Footnote 6 of Fischbacher et al. (2001) that the standard (unique) Nash-Equilibrium of not contributing anything requires common knowledge of rationality, which -in light of a substantial fraction of conditional cooperators in previous PGG experiments- we do not want to assume. Hence, receivers only state their unconditional contribution (which is known to the proposers).

stage. To do so, we use Raven's Matrices (Basteck & Mantovani, 2018). The Raven test is a leading non-verbal measure of analytic intelligence and test scores associated with the degree of sophistication in the beauty contest, with performance in Bayesian updating, and with more accurate beliefs. Within 5 minutes, participants are asked to complete Raven's Matrices, being scored on the number of correct answers minus the number of incorrect answers.

- **Loss aversion:** We use the (incentivized) loss aversion measure in risky decisions from Gächter, Johnson, and Herrmann (2010). High degrees of loss aversion may make participants less likely to adjust their preferences, as they may feel attached to their current partner (*endowment effect*). Although unlikely given the information sets of participants in our experiment, (expectation-based) loss aversion may also influence initial reporting strategies (Meisner & von Wangenheim, 2020). Hence, we elicit an incentivized measure of loss aversion.
- **Socio-demographic controls:** Before concluding the experiment, participants complete a short questionnaire including gender, field of study, final high-school (math) grade and a question on previous experience in economic experiments.

Repetitions We repeat the main stages of the experiment 4 times. This means that participants both repeat the team formation process as well as the PGG 4 times. We implement a "perfect stranger matching" at the group level. While groups remain fixed, every group interacts only once with each other. After each round of team-formation, participants play the PGG with each other without receiving any feedback after submitting their unconditional (and for the proposers also their conditional) contributions.

We elicit the beliefs about the matched partner's contributions only after all 4 rounds are played. We do not announce the belief elicitation before. This rules out that beliefs about the ability to judge the behavior of another player influence (changes in the) preference lists.

Naturally, participants only fill out the initial questionnaires once before the first round starts and administer the cognitive ability test as well as the socio-demographic questionnaire only once after the last round has finished and beliefs have been elicited.

Payoffs and incentive compatibility We choose one round of the PGG to be payoff relevant. Participants earn money both through their final payoff from the PGG (determined by their own and the partner's contribution choice) as well as their point belief about the contributions of their partner. For a correct guess of the unconditional contribution, we pay out a fixed sum. If the guess does not match the actual contribution, participants do not receive any compensation. For the proposers, we randomize whether their conditional or their unconditional contribution is implemented.

In addition to the show-up fee, we also incentivize decisions in the Raven's matrices task and in the loss aversion elicitation.

Through the compensation in the PGG we indirectly incentivize the submission of truthful rank-ordered lists. To have incentive-compatibility for both of the proposers' lists (original and potentially revised one), we randomly choose one of the two to be relevant for determining the final matches. Informing proposers about this is crucial to guarantee that the initial ranking is not perceived as meaningless, given the repeated nature of the experiment.

Setting and sample size The experiment takes place at the Max Planck ECONLAB. We aim for a total sample size of 320 participants (with matching groups of size 8), equally sized across treatments. To facilitate repetitions as planned, this requires having sessions of 32 participants. Sample and matching group size are based on power calculations, outlined after the description of our planned analyses. Due to the ongoing COVID-19 pandemic and associated restrictions to administering regular lab sessions, we resort to an online experiment with the standard subject pool of the laboratory.

3 Hypotheses

3.1 Stability

Our first set of main outcome variables deal with the effect of reciprocating preferences on matching outcomes under the DA mechanisms. We investigate stability of the DA and analyze the individual behavior that may give rise to instability. Our first main hypothesis posits that the fraction of matching groups where the rematching outcome changes after the rematching stage is larger under *Info*.

Hypothesis 1: *Higher instability under Info than under No-Info.*

Our main hypothesis is a direct consequence of the individual behavior of participants we hypothesize to observe. We expect proposers to change their individual preferences more often under *Info* than under *No-Info* which in turn results in higher instability. We further hypothesize that these preference changes are consistent with our underlying theory of reciprocating preferences (Opitz & Schwaiger, 2021). Thus, we hypothesize that preference changes under *Info* are not a result of receiving additional information and adjusting preference lists idiosyncratically (e.g. due to a misperception that not reporting truthfully may result in a better outcome), but systematically consistent with agents having reciprocating preferences. In a nutshell, this means that being ranked favorably by someone else (weakly) increases the likeability of the other. To the contrary, being ranked low by someone (weakly) decreases the likeability of the other. We summarize these auxiliary hypotheses.

Hypothesis 2: *More preference adjustments under Info than under No-Info*

Hypothesis 3: *Participants under Info adjust their preference orders consistent with agents having reciprocating preferences.*

3.2 Public Goods Game

The public good game incentivizes the choice of partners in the team-formation stage. It also allows us to shed light on mechanisms that may drive preference adjustments in the rematching stages. Most importantly, it allows us to understand whether preference adjustments (or null results) are driven by beliefs about receivers contributions and/or whether people show an intrinsic preference for someone who likes them, which may lead to higher conditional contributions, our main hypothesis for the PGG.

Hypothesis 4: *Unconditional contributions of proposers are higher when they interact with a receiver who ranked them high on their preference list.*

Our auxiliary hypotheses investigate the underlying mechanisms of finding (null) effects for unconditional contributions. From the perspective of proposers, we can discriminate between belief-based and preference-based explanations. Proposers may expect differential contributions based on whether the receiver ranked them high or low on their preference lists and adjust their own (unconditional) contributions (e.g. if they are "conditional cooperators"). Alternatively, being highly ranked may result in higher conditional contributions, irrespective of beliefs if underlying social preferences change.

Hypothesis 5: *Proposers expect higher (unconditional) contributions from receivers who ranked them high on their preference list.*

Hypothesis 6: *Conditional contributions of proposers are higher when they interact with a receiver who ranked them high on their preference list.*

Our data also allows us to check for the accuracy of proposers' beliefs by analyzing the unconditional contributions of receivers. We hypothesize that receivers (unconditionally) contribute more to the PGG when they are matched with someone they rank high on their preference lists.

Hypothesis 7: *The unconditional contribution of receivers is higher when they interact with someone they ranked high on their preference list.*

4 Analysis

4.1 Hypothesis Tests

4.1.1 Stability

Hypothesis 1: To test Hypothesis 1, we compare the share of unchanged matching outcomes across conditions (*Info* vs. *No-Info*). We create a binary variable per *matching group*, indicating whether the initial outcome remains constant after the rematching stage. We compare the frequency of changes with a (non-parametric) two-sample χ^2-test.[3]

Hypothesis 2: To test Hypothesis 2, we compare the share of preference changes on the individual level across experimental conditions (*Info* vs. *No-Info*). We create a binary variable per individual proposer (in each round), indicating whether their initial preference order was adjusted in the rematching stage.

First, we compare the frequency of changes with a (non-parametric) two-sample χ^2 test. Second, we run logit models with the binary indicator as an outcome variable. We run the logit regression both solely with a treatment dummy, as well as with our individual controls (gender, cognitive ability, loss aversion). Standard errors are clustered at the individual level.

Hypothesis 3: To test whether participants in *Info* adjust their preference orders consistent with a model of reciprocating preferences (Opitz & Schwaiger, 2021), we first define a binary indicator to classify every preference adjustment. A preference adjustment by a proposer is called *inconsistent* if the preference order changes, but the (now) more favorably ranked receiver did not give a strictly better rank to the proposer compared to the (now) less favorably ranked receiver.

We then describe to which extent preference changes are consistent with a model of reciprocating preferences in *Info* and compare the shares of *consistent* adjustments out of all adjustments across treatment descriptively. With a two-sample χ^2-test, we compare the fraction of *consistent* adjustments out of all decisions across both conditions. Additionally, we run a logit regression with a binary indicator as an outcome variable, taking the value of 1 if a preference adjustment was consistent and 0 otherwise (either inconsistent adjustment or no adjustment). We regress this on a treatment dummy with and without our individual controls (gender, cognitive ability, loss aversion) to corroborate the result from the χ^2-test. Standard errors are clustered at the individual level.

[3]Comparing matching outcomes before and after the rematching stage is a conceptually slightly different approach than deriving stability from the original matching and the (potentially) revised preferences. While those two approaches are equivalent if our hypotheses on individual adjustments hold, there are specific preference adjustments for which the equivalence does not hold. If it is the case in our data, we rerun the analysis and quantify the differences.

4.1.2 Public Goods Game

Hypothesis 4: To test Hypothesis 4, we resort to multiple linear regressions, estimated via OLSs. We can compare proposers who were ranked high and proposers who were ranked low by their partner (Test 7a; within *Info*). We can also hold constant the actual rank a proposer got, and compare beliefs under the situation that proposers knew about this (*Info*) to the situation where proposers were not aware of it (*No-Info*). Hence, we estimate the effect of *knowing* the rank on beliefs by regressing proposer's beliefs on the interaction between the received rank and the treatment (Test 7b; between *Info* and *No-Info*).

Test 4a; Regression equation: $c_{pt} = \beta_0 + \beta_1 * \tau_{pt}(r) + \beta_2 * \tau_{rt}(p) + \beta_3 * X_p + \beta_4 * Z_p + \beta_5 * t + \epsilon_{pt}$

- Comparison of unconditional contributions (c_{pt}) of proposer (p) conditional on rank received by a receiver (r) within *Info*.
- We regress the contributions in each round t on a constant, a linear time trend, time-invariant characteristics of the individual, the rank received by the matched receiver r as well as the rank of the matched receiver in the preference list of proposer p. Our main coefficient of interest is β_2, estimating the causal effect of being ranked higher on the ordered preference list of receiver r on the (unconditional) PGG contribution of proposer p. The coefficient β_1 allows us to compare magnitudes, both within the estimation equation and in comparison to hypothesis 7, estimating the analogous coefficient for the receivers.
- X_p are observed time-invariant heterogeneities across the proposers (gender/cognitive ability/loss aversion).
- Z_p are unobserved time-invariant heterogeneities across the proposers. We cluster the standard errors at the individual level to account for correlations in the error term.

Test 4b; Regression equation: $c_{pt} = \beta_0 + \beta_1 * \tau_{pt}(r) + \beta_2 * \tau_{rt}(p) + \beta_3 * X_p + \beta_4 * Z_p + \beta_5 * t + \beta_6 * Info_p + \beta_7 * [Info_p * \tau_{rt}(p)] + \epsilon_{pt}$

- We fully interact the rank received by a receiver ($\tau_{rt}(p)$) with the treatment status. With the interaction effect (β_7), we estimate the effect of *knowing* one's own rank in the ordered preference list of the receiver on the unconditional contributions, holding the true rank constant.

Hypothesis 5: The test of this hypothesis follows the exact same logic and estimation strategy as compared to Hypothesis 4.

Test 5a; Regression equation: $b(c_{rt})_{pt} = \beta_0 + \beta_1 * \tau_{rt}(p) + \beta_2 * X_p + \beta_3 * Z_p + \beta_4 * t + \epsilon_{pt}$

Test 5b; Regression equation: $b(c_{rt})_{pt} = \beta_0 + \beta_1 * \tau_{rt}(p) + \beta_2 * X_p + \beta_3 * Z_p + \beta_4 * t + \beta_5 * Info_p + \beta_6 * [Info_p * \tau_{rt}(p)] + \epsilon_{pt}$

Hypothesis 6 Again, the test of this hypothesis follows the exact same logic and estimation strategy as compared to Hypothesis 4.

Test 6a; Regression equation: $cc_{pti} = \beta_0 + \beta_1 * c_{rti} + \beta_2 * \tau_{pt}(r) + \beta_3 * \tau_{rt}(p) + \beta_4 * X_p + \beta_5 * Z_p + \beta_6 * t + \epsilon_{pt}$

- i denotes the specific row in the conditional contribution table.

Test 6b; Regression equation: $cc_{pti} = \beta_0 + \beta_1 * c_{rti} + \beta_2 * \tau_{pt}(r) + \beta_3 * \tau_{rt}(p) + \beta_4 * X_p + \beta_5 * Z_p + \beta_6 * t + \beta_7 * Info_p + \beta_8 * [Info_p * \tau_{rt}(p)] + \epsilon_{pt}$

Hypothesis 7: As the decision environment for receivers is identical in both treatments, we pool the treatments and analyze receiver behavior jointly. The estimation strategy is analogous to the other analyses.

Test 4; Regression equation: $c_{rt} = \beta_0 + \beta_1 * \tau_{rt}(p) + \beta_3 * X_r + \beta_4 * Z_r + \beta_5 * t + \epsilon_{rt}$

4.2 Additional Analyses & Robustness Checks

- To further scrutinize what determines individual decisions to adjust their preferences (as a proposer), we run separate logit regressions for each treatment. We include individual controls, as well as the rank of the matched receiver in the proposer's initial preference list. In the *Info* treatment, we include the rank the proposer was assigned by the matched receiver, as well as the proposer's position in the preference lists of the other receivers. We do this both for the binary variable capturing preference adjustments (see Hypothesis 2), as well as for the variable capturing *consistent* preference adjustments (see Hypothesis 3).
- We examine the beliefs of receivers about a proposer's contribution in the PGG depending on the rank given to the proposer. Finding a positive effect for both contributions and beliefs would point to a belief-based explanation [i.e., receivers give more because they expect receivers to give more (e.g. if they are "conditional contributors")], whereas a positive effect for the contributions and none for the beliefs point to a preference-based explanation where receivers are more willing to forgo an individual benefit for maximizing group surplus when working with a proposer they like.
- While our main hypotheses put emphasis on the stability of matching mechanisms and the underlying reasons for instability, we can also analyze potential efficiency losses of having a DA in place in situations where agents with reciprocating preferences have no possibility to opt out of the mechanism and rematch. By comparing cooperation behavior in the PGG of agents within *Info* who were matched according to their initial ranking and those who were matched

according to their revised preferences list. This allows us to analyze the *average efficiency gain (loss)* through being able to incorporate reciprocating preference into the decision.

- For our hypotheses tests 5a/5b, we may run a robustness checks in which we include the information how much the receiver was liked by the proposer in the initial preferences list ($\tau_{pt}(r)$) as a control variable.

4.3 Power analysis

We proceed in different steps to determine our sample size.

- First, we simulate preferences of proposers and receivers. Here, we have to make an assumption on the degree of correlation between those preferences.
- Second, we simulate the resulting matchings based on those original preferences.
- Third, we assume that proposers have reciprocating preferences. Here, we have to make an assumption on the strength of the reciprocating preferences, i.e., how important they are for agents compared to their initial ranking.
- Fourth, we simulate the resulting matchings based on those potentially revised preferences.
- Fifth, we analyze the fraction of stable matchings. Those are the cases where the matchings are the same under the original and the reciprocating preferences.
- Sixth, we determine the required sample size to detect those adjustments between treatments.

We perform all steps above for different group sizes to quantify the trade-off between a higher number of required subjects per observation (stability on the group level) and higher instability through more individual preference changes in the matching group. This means that we base our sample size calculations on testing Hypothesis 1. We allow for a small correlation (0.1) in the initial preferences of the participants and assume that reciprocating preferences are of modest relevance. This means, that we only expect preference changes in cases where two receivers were initially viewed as rather similar. Specifically, we assume that the average importance of being ranked first by someone is one fifth of the importance of the initial assessment. From those assumptions, we derive the fraction of unstable matchings we expect to see in the *Info* treatment and calculate the sample size with which we can identify the hypothesized emergence unstable matches. With a sample of 320 subjects (40 matching groups* 8 participants per group), we are able to detect effect sizes larger than 17 percentage points compared to the baseline *No-Info*, assuming that 10% of the matchings in the baseline are unstable due to idiosyncratic (and treatment independent effects).

With 160 participants per Treatment, we have 40 (320 participants/8 participants per group) observations to analyze stability at the matching group level per round. With four rounds, the number of observations is 160 (4*40). For the analysis of individual preference changes, we analyze all proposers in all rounds (40*4*4 = 640). For the analysis of the 2-person PGG, we have in total 640 dyads for the analysis and 1280 contributions as well as belief elicitations.

4.4 Design Overview

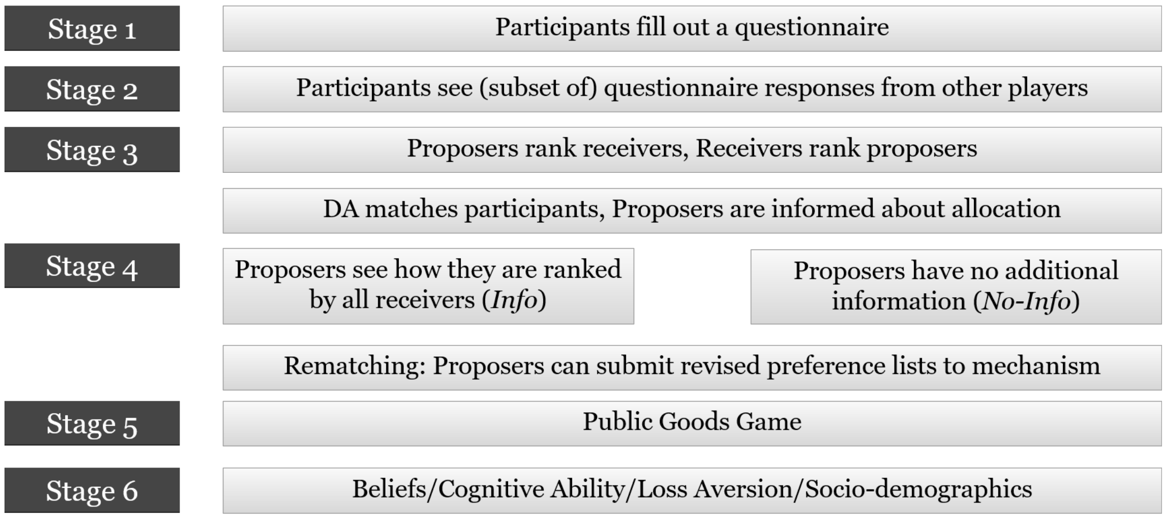

Figure 1: Overview

4.5 Transition from online experiment to laboratory experiment [added: 23 March 2022]

We conducted pilot sessions in June 2021, following the exact design described above. Online implementation proved to be impractical in an experiment with simultaneous interactions of 32 subjects. Participants left the session due to connectivity problems and lack of attention was evident due to resulting waiting times, so it was decided to postpone the main data collection until laboratory experiments were feasible again.

The offline implementation requires only minimal deviations from the original design. These changes concern the repetitions of the main stages (see ”Repetitions” in Section 2). As sessions can be conducted with at most 24 participants simultaneously due to space restrictions, we opt for 5 repetitions of the main stages to ensure power. This requires us to deviate from a ”perfect stranger matching” at the group level and re-shuffle groups randomly (holding the proposer/receiver-roles fixed).

Laboratory experiments are scheduled for March/April 2022 at MELESSA (LMU Munich). We aim for a total sample size of 336 participants.

References

Basteck, C., & Mantovani, M. (2018). Cognitive ability and games of school choice. *Games and Economic Behavior*, *109*, 156–183.

Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, *71*(3), 397–404.

Gächter, S., Johnson, E. J., & Herrmann, A. (2010). Individual-level loss aversion in riskless and risky choices. *CeDEx Discussion Paper No. 2010-20*.

Gächter, S., Kölle, F., & Quercia, S. (2017). Reciprocity and the tragedies of maintaining and providing the commons. *Nature Human Behaviour*, *1*(9), 650–656.

Gale, D., & Shapley, L. S. (1962). College admissions and the stability of marriage. *American Mathematical Monthly*, *69*(1), 9–15.

Goeree, J. K., Holt, C. A., & Laury, S. K. (2002). Private costs and public benefits: Unraveling the effects of altruism and noisy behavior. *Journal of Public Economics*, *83*(2), 255–276.

Hakimov, R., & Kübler, D. (2020). Experiments on centralized school choice and college admissions: a survey. *Experimental Economics*, 1–55.

Haruvy, E., & Ünver, M. U. (2007). Equilibrium selection and the role of information in repeated matching markets. *Economics Letters*, *94*(2), 284–289.

Meisner, V., & von Wangenheim, J. (2020). School choice and loss aversion. *Working Paper*.

Opitz, T., & Schwaiger, C. (2021). Reciprocating preferences in two-sided-matching. *mimeo*.

Pais, J., & Pintér, Á. (2008). School choice and information: An experimental study on matching mechanisms. *Games and Economic Behavior*, *64*(1), 303–328.

Pais, J., Pintér, Á., & Veszteg, R. F. (2011). College admissions and the role of information: An experimental study. *International Economic Review*, *52*(3), 713–737.

Pais, J., Pintér, A., & Veszteg, R. F. (2012). Decentralized matching markets: A laboratory experiment. *Working Papers Department of Economics 2012/08, ISEG - Lisbon School of Economics and Management*.

Spraggon, J., & Oxoby, R. J. (2009). An experimental investigation of endowment source heterogeneity in two-person public good games. *Economics Letters*, *104*(2), 102–105.