Preferences-dependent Learning in the Centipede Game: the Persistence of Mistrust

Astrid Gamba [a*] Tobias Regner [b]

[a] *Universitá degli Studi dell'Insubria, Varese, Italy*

[b] *University of Jena, Jena, Germany*

October 17, 2019

This is the post-print version of an article to be published in the European Economic Review.

Abstract

A candidate explanation for the persistence of heterogeneous behavior in a sequential social dilemma played many times is the existence of heterogeneous preferences. Preferences-dependent conjectures about opponents' behavior are an additional source of heterogeneity. By behaving differently, different preference types acquire different information. Thus, when observing only outcomes of own past interactions heterogeneous and possibly wrong conjectures about opponents' strategies may endogenously arise and persist. In a Centipede game experiment played for forty rounds, we manipulate the type of ex post information and the method of play. We find that, when the game is played in its reduced normal form and subjects have only access to personal statistics, heterogeneity of behavior across preference types persists in the long run. In this case, behavior resembles a self-confirming equilibrium: selfish subjects take at earlier nodes due to an unjustified lack of trust. When subjects have also access to public statistics, heterogeneity disappears: selfish subjects tend to pass more often and play moves towards Bayes Nash equilibrium.

JEL classifications: C72, C73, C9231, D83

Keywords: social preferences, learning, self-confirming equilibrium, experiment

*Corresponding author (Department of Economics, Universitá degli studi dell'Insubria, Via Monte Generoso 71, 21100, Varese, Italy. Phone: +39 3454871374. E-mail: astrid.gamba@gmail.com).

1 Introduction

In many strategic contexts, heterogeneous preferences are a candidate explanation for heterogeneous equilibrium behavior. Besides having a direct effect on behavior, preferences may also have an indirect effect through conjectures: preferences drive behavior which, in turn, determines what agents observe and learn about opponents' strategies. Thus, when agents learn from their own experiences they may form different and possibly incorrect conjectures, resulting in an additional source of heterogeneity. This mechanism is particularly relevant in strategic contexts where agents move sequentially and observe after playing only the outcomes of their own interactions. Stable outcomes where some agents hold incorrect conjectures about the opponents' behavior at off-the-path decision nodes can result.

The Centipede game (henceforth, CG) is a typical example of a strategic situation where heterogeneous behavior is observed, both in initial and in final rounds of recurrent interactions. This game has attracted considerable attention from the experimental literature due to the failure of the backward induction prediction.[1] A strand of the literature that analyzes multiple rounds CG (McKelvey and Palfrey, 1992, Fey et al., 1996, Nagel and Tang, 1998) finds that behavior does not converge to the unraveling prediction—i.e., subjects keep taking on average at middle nodes—and learning mostly takes place in the beginning.[2]

In this paper we provide an explanation for why, even after many recurrent interactions, stable configurations where a majority of subjects take at middle nodes and a minority at late nodes (or never) can emerge. While the focus of most papers on the CG is to explain deviations from the unraveling prediction, the aim of our paper is to explain the source of last rounds heterogeneous behavior. The experimental literature that has specifically studied the sources of behavioral heterogeneity in the CG focused on initial behavior (Kawagoe and Takizawa, 2012, Garcia-Pola et al., 2016, Healy, 2016). Our paper is the first to study the *persistence* of such heterogeneity.

We study experimentally whether and how preferences-dependent conjectures about the opponent's strategy contribute to explain the persistence of heterogeneous behavior in the CG. Looking at stable outcomes of recurrent interactions, we ask whether the exogenous provision of aggregate statistics about past play of subjects in the opponent's role decreases heterogeneity of behavior across preference types. In order to capture the possibility that agents hold incorrect conjectures about the

[1] For a comprehensive review of the literature on the CG see Garcia-Pola et al. (2016).

[2] Nagel and Tang (1998) report interesting patterns of round to round behavior: in initial rounds subjects tend to adapt to previous round observations of the terminal node—i.e., they stop earlier (later, respectively) when the opponent stopped earlier (later, respectively)—but changes diminish over time.

opponent's behavior at off-the-path decision nodes, we adopt the solution concept of self-confirming equilibrium (Battigalli, 1987, Fudenberg and Levine, 1993a). We show that when subjects rely only on their personal statistics about play of subjects in the opponent's role, heterogeneity can persist in the long run, and play resembles a self-confirming equilibrium (henceforth, SCE) with heterogeneous preferences and beliefs. When subjects have access to public statistics about past play of subjects in the opponent's role, heterogeneity disappears and play moves towards a Bayesian Nash equilibrium (henceforth, BNE).

Our intuition is the following. In the CG, as in other Trust-like games, at any decision node (rational) agents decide to pass either because they have preferences for efficiency or because they expect with sufficiently large probability that the opponent will also pass at the next node (or both). Absent intrinsic motivations, selfish agents' incentive to pass is purely beliefs-based. When they have low expectations about the opponent's cooperation at late nodes, they take early and, due to sequential play, they preclude themselves from learning whether the opponent is actually cooperative or not. Thus, when they have only access to their personal observations, their lack of trust can persist in the long run so that they keep taking early, even when it is sub-optimal.[3] Instead, when they have access to aggregate statistics about past behavior of agents in the opponent's role and the level of cooperation at late nodes is high enough, their behavior becomes much closer to those of pro-social types (except at the last node) and thus, heterogeneity will diminish.

In our study, we elicited social preferences online through the Social Value Orientation (Murphy et al., 2011) one week before the experiment. This allows us to classify subjects in pro-selfs—i.e., more oriented towards own payoff maximization—and pro-socials—i.e., more oriented towards social efficiency. We set up an increasing sum payoff structure for the CG that provides high incentives to pass—as the private cost of increasing social efficiency is rather small. We vary both the type of ex post information about behavior of subjects in the opponent's role (Personal vs. Public) and the method of play (Direct vs. Strategy). In Personal, after each round of the CG, subjects observe only the outcomes of their own interaction, while in Public subjects have also access to aggregate statistics about past round behavior of all subjects in the opponent's role. The purpose of this manipulation is to vary the dependence of information feedback on personal past play. For what regards the method of play, in Direct players move sequentially, while in Strategy players

[3]The evolution of preferences literature has considered the possibility that in a dynamic game like the CG (Gamba, 2013) or the Trust game (Adriani and Sonderegger, 2015) incorrect beliefs of selfish types about off-the-path decision nodes prevent them from invading a population of altruists. This intuition is supported also by empirical evidence provided by Butler et al. (2015) who study the relation between individual trust and economic performance.

choose their stopping node simultaneously. Manipulating the method of play has implications on the content of public information feedback. In Public-Direct subjects are told the previous round distribution of *actions* actually played by subjects in the opponent's role at nodes that have been reached with positive frequency; in Public-Strategy subjects are told the previous round distribution of choices over stopping nodes (i.e., *strategies* of the reduced strategic form CG) made by subjects in the opponent's role.

In this setup, we are interested in how long run outcomes vary with the type of information feedback. Since the two methods of play have been found to produce similar behavioral patterns (Nagel and Tang, 1998, Kawagoe and Takizawa, 2012, Garcia-Pola et al., 2016), we expect no substantial differences in behavior in the baseline information treatment. Our first hypothesis is that in Personal, for both methods of play, last rounds behavior is heterogeneous across social preference types, with pro-selfs taking at earlier nodes than pro-socials. We ask whether such outcome is induced by type-dependent conjectures, so that it can be described by a SCE with heterogeneous beliefs. We then test whether this heterogeneity disappears in Public and ask whether this is determined by pro-selfs taking at later nodes due to more accurate beliefs. Notice that in Public-Direct, due to sequential play, the feedback that a subject receives depends on past play of the other subjects with his role,[4] and thus information provision may not solve subject's uncertainty about the objective distribution of strategies played by the opponents. Instead, in Public-Strategy, subjects are informed about the frequencies of strategies and play should move towards BNE.[5]

We find that last rounds behavior differs significantly across social preference types only in Personal-Strategy—and not in Personal-Direct—where pro-selfs take at earlier nodes than pro-socials. Focusing on Strategy, public information provision reduces last rounds heterogeneity of behavior across types, due to the increase in pro-selfs' passing rates. We show that while in Personal-Strategy initial heterogeneity across preference types persists along the learning process, in Public-Strategy pro-selfs adapt behavior to exogenous feedback and they soon close the gap with pro-socials, so that heterogeneity disappears. Looking at beliefs, we do find that pro-selfs underestimate average passing rates at middle-late decision nodes of the opponent in Personal-Strategy, while they are rather accurate in Public-Strategy. Interestingly, pro-self more than pro-socials tend to adapt behavior to observations

[4]In the extreme case in which a node of the opponent is not reached in any match in the previous round, no observation is made and no feedback is provided.

[5]While in last rounds of Public-Strategy, beliefs about strategies should be common and correct, in Public-Direct they may not be correct, so that play should be better described by a SCE with unitary (i.e., common) beliefs (see Fudenberg and Levine, 1993a).

and aggregate information affects passing rates in Strategy but not in Direct.

A strand of the literature on the CG considers the possibility that heterogeneous behavior is driven by heterogeneous preferences. McKelvey and Palfrey (1992) interpret the failure of the unraveling prediction through a model of incomplete information, whereby few altruists always pass and selfish types mimic them to earn more. Yet, with the exception of Healy (2016), who finds support for a model of incomplete information with rational players and heterogeneous preferences, subsequent experimental work on the CG disregards social preferences as the main determinant of initial heterogeneity (Fey et al., 1996, Garcia-Pola et al., 2016).[6] Evidence is rather in favor of models that weaken rationality (quantal response equilibrium) or common knowledge of rationality (level-k thinking).[7] Our paper complements this literature by showing that *long run* heterogeneity can be explained by heterogeneous preferences and, in particular, that behavioral heterogeneity across preference types can emerge as an equilibrium phenomenon, due to type-dependent learning. Indeed, when public feedback closes the information gap between preference types, long run heterogeneity disappears.

Our paper departs from this literature in that it interprets (last rounds) behavioral heterogeneity without questioning rationality or common knowledge of rationality, but only weakening the correctness of beliefs requirement that characterizes Nash equilibrium. Similarly, recent contributions on the CG explain the observed failure of the unraveling predictions through incorrect equilibrium beliefs. Cox and James (2012) suggest that uncertainty about the opponent's node-specific play might explain the absence of unraveling and point at analogy based expectation equilibrium (Jehiel, 2005) as a possible tool to interpret observed behavior.[8] Danz et al. (2016) test experimentally whether in a CG subjects' beliefs actually bundle opponent's information sets into analogy classes by manipulating information provision about the opponent's node-specific behavior. They find that players make actually good use of their own past experiences at specific information sets and exogenous provision of public information crowds in the use of private statistics.[9]

[6]In particular, Garcia-Pola et al. (2016), through a systematic experimental test of all the main competing models of behavior in the CG, conclude that preferences-dependent explanations are dominated by bounded rationality models (quantal response) and models based on the failure of common knowledge of rationality (level-k reasoning). Also, Embrey et al. (2017) find that the existence of cooperative types have limited effect on the degree of cooperation in a repeated Prisoner Dilemma.

[7]Kawagoe and Takizawa (2012) find that a level-k thinking model explains data better than an agent-quantal response equilibrium model.

[8]Differently from quantal response equilbrium and level-k reasoning, both SCE and analogy based expectation equilibrium do not challenge the rationality assumption. The main difference between the two notions is that the latter imposes that incorrect conjectures about node-specific behavior have a particular structure (analogy classes).

[9]Mermer and Suetens (2017) study experimentally whether uncertainty about the second-

It is worth stressing that node-specific incorrect beliefs in this literature actually imply that play does not unravel, the opposite of our prediction (and findings). What to expect about the effect of public information release on behavior in the CG naturally depends on the expected passing rates. We derive our behavioral predictions under the assumption that passing rates are sufficiently high, which is justified by the payoff structure that we implement. As a consequence, in our setting, resolving for subjects' uncertainty about the distribution of strategies played by the opponents implies more cooperation by pro-selfs and thus less unraveling. It has been argued in the literature on the CG that not only the failure of the unraveling prediction in general, but also the effect of social preferences on behavior, are very sensitive to the payoff structure implemented (see, for example, Fey et al., 1996). Moreover, Maniadis (2012) finds that aggregate information release produces different effects on passing rates depending on the payoff structure.[10]

A link between incorrect beliefs and social preferences like the one we provide in this paper is still missing in the literature on the CG. The novel feature of our design is that it allows us to analyze how varying information feedback affects long run behavior of different preference types. The first contribution of our paper is that it provides an explanation for the persistence of behavioral heterogeneity in the CG based on social preferences and preferences-dependent conjectures. Yet, the mechanism that we describe goes beyond the CG and it can apply to other dynamic games played recurrently where social preferences and sequential rationality matter (e.g., trust-like games). What is crucial is that the ex post information structure is such that some strategies prevents a player from observing the opponent's strategy after playing, which implies that the extent of strategic uncertainty can be preferences-dependent.

This leads to our second key contribution as the characteristics of the equilibrium configurations in a dynamic social dilemma may depend on the type of information feedback. Typically, if subjects can observe only outcomes of own past interactions, off-path prediction errors can persist in the long run. In this case, aggregate behavior can be described by a solution concept like SCE that contemplates that conjectures about the opponent's strategies are heterogeneous and possibly incorrect. When players have access to detailed public statistics, aggregate behavior resembles BNE, where all preference types have common and correct beliefs. This contribution has a methodological implication: by varying the ex post information structure of exper-

mover's type in a CG can increase the first-mover's performance in terms of material payoffs.

[10]Embrey et al. (2017), through a meta-analysis of the repeated Prisoner Dilemma, which shares with the CG the tension between social efficiency and sequential rationality, conclude that payoffs of the stage game (among other parameters) affect initial play, by acting on the value of cooperation.

imental games, experimenters can impose convergence to a particular equilibrium. This could be a serious concern, as experimenters may rationalize stable outcomes with a notion of equilibrium, which is the result of a design choice.

The paper is organized as follows. In the next section we provide the theoretical framework and the learning interpretation. Section 3 describes design, hypotheses and procedures of the experiment; in Section 4 we present the experimental results; Section 5 concludes.

2 Self-confirming equilibrium and the Centipede game

In the experiment, we will employ the Centipede game (CG) displayed in Figure 1. We will derive our experimental hypotheses about last rounds behavior from the assumption that subjects play according to a self-confirming equilibrium (SCE). Let us first define the notion of SCE that we adopt to interpret experimental data and then the learning process that can deliver such SCE in the last rounds.

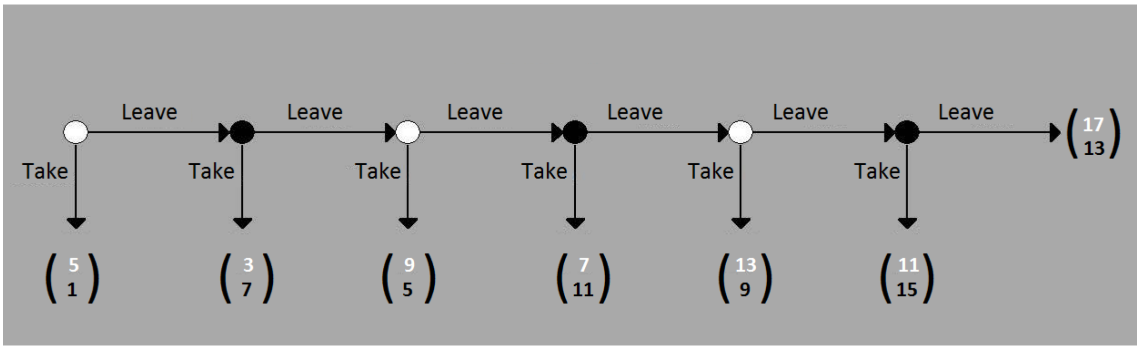

Figure 1: The Centipede game

Assume that for each player (i.e., role) $i = \{White, Black\}$ there is a large population of agents with heterogeneous preferences. Each agent has a preference type $\theta \in [0,1]$. Denote $q_i(\theta)$ the share of preference type θ in population i. Agents are drawn at random to play the stage game and each agent plays a pure strategy $s_{i,\theta} \in S_i$. Hence, each player (role) i plays a mixed strategy $\sigma_i \in \Delta(S_i)$, induced by q_i and the pure strategies adopted by each preference type in i's role. We allow agents in population i to have heterogeneous conjectures on the opponent j's (mixed) strategy: $\mu_{i,\theta} \in \Delta(S_j)$. Assume that agents do not know the distribution of preference types in either population. Denote $\pi(z|s_{i,\theta};\sigma_j)$ the objective probability that preference type θ in population i observes terminal node z given his own move and the mixed strategy of the opponent. Denote $\rho(z|s_{i,\theta};\mu_{i,\theta})$ the subjective probability of observing terminal node z as assessed by preference type θ in population i given his own strategy and his conjecture about the opponent's mixed strategy. Assume that after playing agents can only observe the terminal node reached in their own match. We adopt the following definition of SCE, which adapts the definition of Dekel et al. (2004), related to static games, to our extensive-form game with heterogeneous preference types.[11]

Definition 1 *A profile of mixed strategies $(\sigma_i)_{i\in I}$ is a* self-confirming equilibrium *if for each preference type θ we can find a conjecture $\mu_{i,\theta}$ s.t. for each $s_{i,\theta} \in supp\ \sigma i$*
i) $s_{i,\theta} \in \arg\max_{s_i \in S_i} \left[\sum_{s_j \in S_j} \mu_{i,\theta}(s_j) U_\theta(s_{i,\theta}, s_j)\right]$ *and*
ii) $\forall z \in Z, \rho(z|s_{i,\theta};\mu_{i,\theta}) = \pi(z|s_{i,\theta};\sigma)$.

The first condition is the standard rationality assumption—i.e., players play best replies to their conjectures about the opponent's strategy; the second condition requires that for each individual the statistical distribution of observations over terminal histories coincides with his subjective probability distribution.

The SCE that we have defined can be seen as the result of a learning process whereby different preference types interact recurrently and anonymously and learn about the opponent's behavior only from their own experiences. Fudenberg and Levine (1993b) provide a steady state learning foundation for SCE for static games.[12] We adapt this interpretation to our extensive form game with heterogeneous agents.

Assume that each population plays recurrently CG. In every round agents are drawn at random and matched to play the stage game with a different opponent. After each play an agent obtains a feedback on the opponent's play and on the basis of all feedback collected in his own matches he updates his beliefs via Bayes' rule. We assume that agents have *ex post perfect recall* (Battigalli et al., 2015), i.e., after playing they remember the opponent's decision nodes that they have reached and the actions played by the opponent at those nodes. In addition, assume that agents believe that opponents are playing according to a stationary distribution, as in a *fictitious play* learning model (Brown, 1951). If there are restrictions to the evidence agents can collect (the feedback is not informative enough), long-run frequencies of personal observations do not unambiguously identify the objective distribution of

[11] In fact, the definition of SCE of Fudenberg and Levine (1993a) is for extensive-form games. However, in their model, differently from Dekel et al. (2004), it is assumed that players observe the terminal histories, which implies that players observe the opponent's type. In this case players would be able to learn the strategies of their opponent. The version of Dekel et al. (2004) is similar to the notion of 'conjectural equilibrium' introduced by Battigalli (1987) and Battigalli and Guaitoli (1997).

[12] On the learning foundation of SCE, see also Fudenberg and Kreps (1995).

the opponent's play. Hence, it can occur that agents hold beliefs that are consistent with the empirical frequencies that they have observed (due to ex post perfect recall) but misrepresent the true distribution (due to uninformative feedback).

Assume that agents have access only to their 'personal' database of experiences, that consists of terminal nodes reached in the games they played. Keeping track of personal experiences allows every agent i with type θ to learn the conditional frequencies of opponent's actions at information sets (decision nodes) visited with positive frequency under $(s_{i,\theta}, \sigma_j)$, a sub-collection of the opponent's information set visited with positive frequency under (σ_i, σ_j). Due to the dependence of the individual strategy on preference type, the information collected is endogenous and depends on the type. Thus, if agents do not experiment enough, the learning process will naturally deliver a SCE with heterogeneous and possibly incorrect beliefs, whereby agents playing in the same role i have different equilibrium beliefs about the opponent j's strategy. Interpreting the strategic interaction as a two-player game (and not as a two-population game), in a SCE with heterogeneous beliefs, for every player i, each pure strategy s_i in the support of the equilibrium mixed strategy σ_i has a different justifying belief.[13]

Alternatively, assume that agents have access also to a 'public' database of terminal nodes reached in all games played by agents in their own population. Keeping track of public statistics of this kind allows all types to learn the conditional frequencies of opponent's actions at information sets visited with positive frequency under (σ_i, σ_j). Thus, in the long run we would have a SCE whereby agents playing in the same role share a common belief about the strategies played in the opponent's population. Such an equilibrium is called a SCE with *unitary* beliefs in the terminology of Fudenberg and Levine (1993a), to distinguish it from the more general notion of SCE that assumes heterogeneous beliefs. Adopting the perspective of a two-player game interaction, in a SCE with unitary beliefs, all pure strategies played with positive probability in equilibrium by every player i are best replies to the *same* belief μ_i.[14] Notice that equilibrium unitary beliefs about the strategy of the opponent do not need to be correct as, even with public information, agents playing in i's role may

[13] Agents' long run beliefs and behavior will depend on their own past behavior. Trivially, if an agent starts the learning process with no trust toward the opponent and keeps stopping the game at the very first node, his beliefs about the opponent's behavior at off-the path nodes will keep being confirmed by the evidence. Hence, he will never learn whether there are agents who actually behave more cooperatively than he expected.

[14] Notice that in games with observable deviators, and hence in two-player games, a SCE with unitary beliefs is equivalent to a Nash equilibrium—see Fudenberg and Levine (1993a), Corollary of Theorem 4 in Section 6, and the related discussion in Kamada (2010). However, the assumption of observable deviators is not satisfied in our context, as players' types are determined by a chance move; hence we cannot use this theoretical result to predict outcome equivalence with Nash in Public-Direct.

not observe under σ_i the off-path intended play of j.

When agents have access to the public database of terminal nodes, what agents playing in role i learn about the opponent j's strategy is exogenous to their own play but not to the play of other agents i, and thus to the distribution of preference types in population i that determines σ_i. Instead, if we assume that agents have access to a 'public' database of *strategies* played on average in the opponent's population, then common beliefs about the opponent's strategy need to be *correct* and long run play would correspond to a BNE.

2.1 An example

Let us explain our intuition with an example for the CG illustrated in Figure 1. Suppose that there are only two types, selfish and altruistic, playing either in the role of White or Black, and that the share of altruists is $q = \frac{1}{3}$. While selfish types maximize their own material payoff, altruists maximize the joint payoff. Any profile of strategies where selfish (altruistic) types play Take (Pass) at every node is a SCE of this game. Such SCE selection is supported by the following system of beliefs: a selfish White attaches a probability larger than $\frac{2}{3}$ to the set of Black's strategies that prescribe Take at Black's first decision node and selfish Blacks attach a probability larger than $\frac{2}{3}$ to the set of White's strategies that prescribe Take at White's second decision node. Obviously, such conjectures are wrong as the probability of observing Take at both nodes is exactly $\frac{2}{3}$, the share of selfish types in the population.[15] Yet, by sticking to playing Take whenever they can selfish players do not learn the true probabilities of the opponent's strategies and continue to erroneously attach such high probabilities to Take. This system of incorrect conjectures delivers a stable outcome that consists of either reaching the first terminal node (with probability $\frac{6}{9}$), or the second (with probability $\frac{2}{9}$) or the efficient one (with probability $\frac{1}{9}$), depending on the preference types matching.

3 The Experiment

Our experiment is set up to analyze whether persistent behavioral heterogeneity in the CG is due to recurrent interaction of heterogeneous preference types with possibly heterogeneous beliefs. To obtain a measure of subjects' social preferences, we used the Social Value Orientation slider measure (SVO, Murphy et al., 2011),

[15]Notice that the conjecture of a selfish Black about the initial action of White must be correct and such that he attaches a probability of exactly $\frac{1}{3}$ to Pass.

via an online platform. In addition, in the first part of the experiment, subjects play a Trust game, whose results we use to benchmark the SVO-based measure (see Section 3.2 for details).

In the second part of the experiment, subjects play for forty rounds the CG represented in Figure 1, which illustrates the game form with monetary payoffs (denoted in experimental currency units, i.e., ECU). The payoff structure displays the usual tension between selfish motives and efficiency concerns, but differently from the original version it provides high incentives to continue at every decision node. As Fey et al. (1996) point out, increasing-sum payoff structures trigger the role of social preferences. We calibrate payoffs so to have a consistent share of subjects playing altruistically at the last node.[16] This makes it optimal even for a own (material) payoff maximizer to reach that node.

In what follows, we first describe the experimental design in details and then the experimental procedures.

3.1 The experimental design

Our experimental design varies the type of information feedback that subjects can obtain about the average behavior of subjects in the opponent's role. In a *between subjects* design we implement a two-fold manipulation.

Manipulation of the ex post information structure. In a first treatment, after playing each round, subjects are informed about the terminal node reached in their own match. We call this treatment Personal, as the information feedback is subject-specific and forms the subject's personal database. In another treatment that we call Public, besides observing own outcomes, subjects also receive aggregate information about *average* behavior of subjects playing in the opponent's role, averaged across all matches occurred in the past round. The purpose of this manipulation is to decrease with aggregate information provision the type-dependence of strategic uncertainty, a possible source of persistence of heterogeneous behavior. Indeed, in Public, the feedback is exogenous with respect to own play, and thus it is independent of own preferences. The content of the aggregate feedback depends on the method of play.

Manipulation of the method of play. We employ two different methods of play. In the first, Direct, players move sequentially in the CG illustrated in Figure 1 until either of the two stops the game by playing Take. In the second, Strategy,

[16] The payoff structure that we implement is similar to that of Healy (2017)—in the "CENT-LO treatment", where it is found that social preferences explain heterogeneous behavior in initial rounds of the CG.

players play a reduced normal form CG where they have to choose simultaneously at which node to stop the game by playing Take, independently of the opponent's decision—as in Nagel and Tang (1998), Kawagoe and Takizawa (2012), Garcia-Pola et al. (2016). Their possible strategies are: Take at the first node (s^1), Pass at the first and Take at the second node (s^2), Pass at the first and at the second node and Take at the last node (s^3), Pass at all nodes (s^4).[17] Subjects knew they had to play one of them and that their final payoffs depended on the choices made by both.

In Personal, where subjects are informed only about the terminal node reached in their own match, ex post information is constant across methods of play.[18] Instead, in Public the aggregate component of the ex post information structure varies with the method of play. In Public-Direct aggregate feedback is node-specific and consists of *conditional frequencies of actions* played in the previous round by subjects in the opponent's role; in Public-Strategy aggregate feedback coincides with *frequencies of the reduced strategies* played in the previous round by subjects in the opponent's role.

Feedback is potentially less informative in Public-Direct than in Public-Strategy. In the former, subjects are only informed about behavior of subjects in the opponent's role at information sets that have been reached with positive frequency in previous round matches. Thus, the node-specific information that a subject can acquire depends on how the other subjects in his role played the game. Indeed, if a node is not reached, no observation is made at this node and no feedback is provided.[19] If a node is reached only in some matches, observed average behavior at this node, and thus the feedback, may not reflect the objective distribution of actions as induced by the underlying distribution of strategies of the opponents. In contrast, in Public-Strategy, subjects are informed about how opponents *intended to play* on average in the past round. Thus, the feedback received by subjects in either role is independent of past choices of the other subjects in their same role.

Notice that only in Public-Strategy aggregate information can entirely solve subjects' uncertainty about the distribution of strategies played by subjects in the op-

[17] Notice that the game structure that we call Strategy is different from the strategy method introduced by Selten (1967). Indeed, we do not ask subjects to make a contingent choice at every information set of the game tree. So, if for example a subject plays strategy 2, we are not informed about how he would behave at the last node in case of crossing it.

[18] While the player who stops later observes ex post the strategy played by the opponent, the player who stops earlier observes only the actions chosen at information sets that he has reached.

[19] When in Public-Direct a subject receives a node-specific feedback about the behavior of subjects in the opponent's role, he might infer that at least one subject in his own role must have reached that node. This information can condition his behavior through imitation. Yet, it is not clear how this effect could interact with preferences. Notice that also the absence of node-specific feedback can condition behavior. However, in the experiment, in every round of Public-Direct, all information sets of the game were reached with positive frequency.

ponent's role. The feedback in this treatment provides them with the relevant information that they would otherwise miss, due to their preferences and behavior—as it may occur in Personal—or to the preferences and behavior of the other subjects in their same role—as it may occur in Public-Direct. Therefore, the two public information treatments can provide interesting insights on how the content of public information feedback (actions versus strategies) affects the learning process, and thus long run outcomes.

The adoption of the two methods of play has implications also for last round beliefs. Given that in Personal, subjects accumulate evidence only about actions actually played by the opponents they have been matched with, *heterogeneous* beliefs across subjects in the same role and off-path prediction errors about the opponent's behavior may persist in equilibrium. In Public-Direct, subjects in the same role should share *common* beliefs about the opponent's actions at information sets that have been reached with positive frequency. Still, they might be incorrect, as they depend on past behavior of subjects in their own role. Instead, in Public-Strategy, given that subjects are informed about the distribution over strategies played by subjects in the opponent's role, their equilibrium beliefs should be *common and correct.*

3.2 Preferences elicitation

We elicited *social preferences* and classified subjects into preferences types in the following way. First, a week before the experiment we administered the SVO test (Murphy et al., 2011) with hypothetical payoffs.[20] Essentially, the SVO test derives an index (measured in angle degrees) based on the relation between the average monetary payoff that a person would allocate to herself and the average monetary payoff that she would allocate to another person (π_o), averaged across a menu of dictator choices (see Appendix A for details).

We use the SVO measure as a proxy for subjects' concern for the payoff of the other. We normalize the SVO angle to get a variable $\theta \in [0, 1]$, that indicates a subject's pro-social attitude and represents the empirical counterpart of the parameter θ in the theoretical framework. The larger θ, the larger the concern for the opponent's payoff. We also use the types classification proposed by Murphy et al. (2011) to categorize subjects into pro-socials and pro-selfs, applying the normalization that we use for the SVO angle also to their proposed threshold. Accordingly, we classify subjects

[20] Mentzakis and Mestelman (2013) compare hypothetical SVO measures to those elicited with salient monetary incentives and find no differences. Greiff et al. (2018) find differences in a within-subjects comparison of incentivized and hypothetical SVO choices but not in a between-subjects comparison.

with a θ above 0.59 as pro-socials (139 subjects), and remaining subjects (117) as pro-selfs, the latter group being formed by 112 individualists and 5 competitive types according to the categorization of Murphy et al. (2011).

Second, the experiment begins with a Trust game played with the strategy method (see Appendix B for details). The strategic context of this game features the same tension between joint payoff maximization and selfish motives that we find in the CG. When the trustor sends money to the trustee the surplus increases but, if the trustee does not reciprocate, the trustor receives less than what he could have received by keeping the money for himself in the first place.

We use choices made in the Trust game to check how meaningful is the SVO-based threshold in this strategic context. We estimate kernel densities for the distribution of θ conditional on choices in the Trust game. These estimates show that when θ crosses the level of 0.64—which closely approximates the normalized equivalent of the threshold of Murphy et al. (2011)—there is a change in trusting and honoring trust behavior (see Appendix B for details). When considering the threshold of 0.64 to classify subjects in pro-selfs and pro-socials, 14 additional subjects are considered as pro-self in comparison to the SVO-based categorization. While we use the standard SVO-based classification of types in our analysis, we will perform robustness checks based on the alternative categorization (see discussion in Section 4.4).

3.3 Beliefs elicitation

We asked for subjects' point beliefs about choices of subjects in the opponent's role before the first round started and in rounds 17, 18, 19 and 40 after decisions were made.[21] In Public, we elicited beliefs before aggregate information about subjects' choices appeared. In Direct, Whites (Blacks) are asked to estimate for each of the three decision nodes of Blacks (Whites) the percentage of Blacks (Whites) playing Pass at that node. This provides us with subjects' expectations about the opponent's choice at the next node. In Strategy, subjects are asked to estimate the percentages of subjects in the other role playing strategies s^1,s^2, s^3 and s^4, respectively. Subjects received 0.5 ECU for each estimate within 5% of the actual value and they knew that they will be informed about their earnings from the estimations at the very end of the experiment.

[21]Although having beliefs data from all rounds would have been ideal, we decided to elicit beliefs only in a subset of all rounds in order not to overload subjects cognitively. Moreover, the elicitation of beliefs was not announced in the instructions and, hence, came at least to some extent as a surprise. We did this to keep subjects' attention focused on the choices in the game. See Appendix C for translated screenshots of the instructions that accompanied the belief elicitation.

3.4 Participants and Procedures

We recruited 256 participants among students from various disciplines at the local university using the ORSEE software (Greiner, 2004). In each session gender composition was approximately balanced and subjects took part only in one session. The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007) and took, on average, 100 minutes.

Upon arrival at the laboratory subjects were randomly assigned to one of the computer terminals. Each computer terminal is in a cubicle that does not allow communication or visual interaction among the participants. Participants were given time to privately read the instructions and were allowed to ask for clarifications. Only after subjects completed part 1 of the experiment, i.e., the Trust game, they received instructions for part 2, i.e., the forty-period CG (see the Appendix for the full set of the experiment's instructions). In order to check the understanding of the instructions subjects were asked to answer a set of control questions. After all subjects had answered the questions correctly the experiment started.

Subjects are randomly assigned to be either White or Black before the first round of the experiment. As told in the instructions, they keep their role for 40 rounds and in every round they are randomly matched in pairs. In order to minimize reputation concerns we used a rotation matching that ensured that during the first 16 rounds subjects cannot affect the decisions of future subjects they will be paired with through their choices in the current match.[22] Subjects were informed about the protocol and its meaning and that their identity was never revealed. This approach allowed us to minimize reputation concerns among subjects in a controlled way.[23]

At the end of the experiment subjects were paid in cash according to their performance. Privacy was guaranteed during the payment phase. As our experiment consists of various components (the Trust game, the forty-period CG and beliefs within it) we had to make a choice about which of them to pay. The majority approach in economics experiments with multiple decisions is to pay for the outcome from every decision made (Azrieli et al., 2018). However, as discussed, for instance, in Charness et al. (2016) and Azrieli et al. (2018), which payment mechanism is best depends on the environment and assumptions about a specific utility theory.[24]

[22] The 32 subjects of one session were divided into two groups, White and Black. Each of the 16 White subjects was matched with each Black subject exactly once during the first 16 rounds. This allowed us to run 16 rounds of play without interaction or contagion effects. In round 17 the protocol started anew as well as in round 33.

[23] Eliminating reputation concerns entirely was not possible in our lab since we decided that studying learning effects required us to play more than 16 rounds (the lab's maximum capacity divided by two).

[24] Note that also the way choices are presented (sequential vs. in a list) may matter for incentive

In Appendix D we provide a discussion about the advantages and drawbacks of the pay-one and pay-all mechanisms in our context and explain the motivation behind the payment mechanism choices we made.

We decided to pay both games and the beliefs, since we regarded the possible wealth effects as minor compared to a potential dilution of incentives to play the 40 round CG, the focus of our experiment. Within the CG we opted to pay two rounds (one from the first, one from the second half) in order to limit the scope for experimentation, that might be a concern in strategic interaction environments with multiple rounds and feedback. Subjects learned about the payoffs from each component at the very end of the experiment: even though this might enhance hedging, getting information about behavior in the Trust game could contaminate behavior and beliefs in the CG.

Thus, the instructions informed subjects that earnings consisted of the payoff from the Trust game in part 1 and the payoff from two randomly chosen rounds of part 2 (one from the first 20, one from the last 20). On top, subjects received earnings from the estimations. The average earnings in the experiment amounted to €27.41 (including a €2.50 show-up fee and €3 for completing the online survey preceding the experiment). Decomposing these earnings into their origin, subjects, on average, received €2.65 from part 1, €8.46 from the first half round, €8.10 from the second half round, and €2.70 from the beliefs elicitation.

3.5 Hypotheses

We are interested in comparing long run outcomes across information treatments.[25] We argue that in Personal, equilibrium configurations with heterogeneous behavior and beliefs, like in a SCE, can emerge. Instead, in Public, long run play should be closer to BNE, with subjects playing best replies to the actual frequencies of strategies played by subjects in the opponent's role. In Public-Direct, subjects playing in the same role should have the same beliefs about the opponent's behavior in the last round—i.e., beliefs should be common among types—while in Public-Strategy, they should be common *and* correct, as in this information treatment subjects can learn the opponents' average strategy.

To start with, thanks to our preferences elicitation, we can test whether in Personal behavioral heterogeneity actually corresponds to heterogeneity across preference types, with pro-selfs taking earlier than pro-socials. Thus, our first hypothesis

compatibility, see Brown and Healy (2018).

[25]We will consider the last ten rounds, that is, when behavior stabilizes.

concerns the heterogeneity of long run behavior across preference types (pro-self vs. pro-social) in the baseline information treatment.

H1 (Persistent heterogeneity). In the last rounds of Personal, behavior and beliefs are heterogeneous across preference types with pro-self subjects taking at earlier nodes than pro-social ones.

Our next step is to disentangle whether last rounds heterogeneity can be entirely explained by heterogeneous preferences (direct effect of preferences on behavior) without the consideration of type-dependent learning (indirect effect of preferences on behavior through conjectures). First, by combining actual passing rates and last rounds beliefs, we will be able to assess whether pro-self types are playing best replies to correct conjectures, like in a BNE, or to wrong conjectures, like in a SCE. In the first case, the answer would be affirmative, i.e., social preferences alone explain last rounds heterogeneity. In the second case, pro-selfs' erroneous conjectures would be a key ingredient and an additional source of persistent heterogeneity. Secondly, if last rounds heterogeneity across preference types in Personal is simply due to heterogeneity of preferences, we should observe the same heterogeneity in Public (ceteris paribus), where the preferences-dependent conjectures component is expected to be absent.

The direct effect of preferences on behavior should be at work even in the initial rounds. Thus, to identify the direct effect of social preferences on behavior, not filtered through the channel of beliefs and experience, we will also investigate whether, given our preferences classification, heterogeneity of behavior across preference types exists from the very first round.[26] We will test whether, given beliefs, pro-selfs take earlier than pro-socials since the very beginning. The literature focusing on initial heterogeneity of behavior in the CG (Kawagoe and Takizawa, 2012, Healy, 2016, Garcia-Pola et al., 2016) finds mixed evidence with respect to a preferences-based explanation. Notice that last rounds heterogeneity of behavior and independence of behavior of preference types in the initial rounds are two compatible pieces of evidence. Our design allows us to shed light on whether the dependence of behavior on preferences is enhanced by learning from own experience and is thus an equilibrium phenomenon.

Even if pro-socials and pro-selfs behave similarly in initial rounds—either because pro-selfs are simply experimenting with passing or because they truly expect that the opponent will also pass—the same observations might induce pro-selfs to unravel and pro-socials to keep passing. Thus, initial heterogeneity is not necessary but it

[26] Notice that analyzing the source of first rounds heterogeneity of behavior serves also the purpose of checking the validity of our preferences classification.

is certainly propedeutical for long run heterogeneity: if pro-selfs tend to take early since the very first rounds, the hypothesized result follows immediately.

Our next experimental hypothesis is derived under the auxiliary assumption—which is a testable hypothesis in itself—that in the last rounds (in all treatments) passing rates are such that it would be rational even for a own payoff maximizer to pass at all nodes except the last one of Black. Specifically, a passing rate larger than $\frac{1}{3}$ at every decision node of the opponent would make it optimal for a own payoff maximizer in Direct to pass at every node, except the last one of Black. It can be easily shown that in Strategy a passing rate larger than $\frac{1}{3}$ is a sufficient condition for a rational own payoff maximizer (with correct beliefs) in White's (respectively, Black's) role to play s^4 (respectively, s^3). If passing rates satisfy this condition, only incorrect conjectures could prevent rational subjects from passing at any information set.

The release of aggregate information reduces strategic uncertainty at nodes that are not reached frequently enough under individual learning. Thus, subjects in either role should end up having common beliefs about the opponent's behavior, provided that they make good use of public information. In particular, pro-self subjects have now access to a richer database of observations. If our auxiliary assumption about passing rates is verified—i.e., in both information treatments at all nodes passing rates are larger than $\frac{1}{3}$—public information should induce pro-selfs to revise their possibly incorrect conjectures and to play Pass more frequently, pushing aggregate behavior towards BNE play.

H2 (Effect of aggregate information). Heterogeneity of long run behavior across preference types decreases from Personal to Public and pro-self subjects take at later nodes.

By design, the content of the public information feedback varies across methods of play. The method of play manipulation has implications on last rounds beliefs and thus on the nature of equilibrium configurations that we should expect (BNE or SCE). As we have discussed in Section 3.1, compared to Personal-Direct, in Public-Direct subjects can derive their predictions from a richer database that includes the observations of other subjects playing in the same role. Yet, such information may not be sufficient to fully correct off-path prediction errors. What subjects in a given role learn about the opponent's strategy still depends on how the other subjects in their role played the game in the past, and thus on their types distribution. Hence, in Public-Direct we expect subjects playing in the same role to have common, but not necessarily correct, beliefs about the opponent's strategy in the last round. As discussed in Section 2, stable situations where different types are playing best replies to the same belief about the opponent's play can be described by a SCE with unitary (i.e., common) beliefs (see Fudenberg and Levine, 1993a).

Instead, in Public-Strategy, subjects receive exogenous information that consists of the *actual* frequencies of strategies played by subjects in the opponent's role. Importantly, such information does not depend on their own strategies nor on strategies played by other subjects in the same role. Thus, subjects' last round beliefs should be common *and* correct, like in a BNE.

To obtain heterogeneity of last round beliefs—with pro-selfs underestimating the opponent's passing rate—it is not necessary that *initial* beliefs are also heterogeneous. Last round heterogeneity of beliefs can be the result of a learning process whereby pro-self and pro-social types adapt behavior to observations in a different way. Given the same conjectures—or the same statistics—about the opponents' behavior, a pro-social subject may decide to Pass, while a pro-self one may decide to Take. Yet, if pro-self subjects hold low expectations about the opponent's passing rate since the very beginning—for the so called false consensus (Ross et al., 1977)—the type-dependence of the learning process is exacerbated.[27] Given that we elicited beliefs in the first, in the middle and in the last rounds, we can test whether type-dependence of beliefs exists ex ante and persists in the long run. Moreover, we can analyze how different preference types react to personal and public information to better characterize how beliefs and behavior evolve and whether their dynamics is type-dependent.

4 Results

We begin by looking at aggregate behavior in the CG. Then, we analyze individual, node-specific choices and beliefs in the last ten rounds in order to test our hypotheses. Finally, we explore the evolution of behavior starting from initial rounds and considering how subjects adapt to different types of information.

4.1 Aggregate behavior

We start with a display of subjects' choices. The histograms in Figure 2 show the distribution of the seven possible nodes that can be reached in Personal as well as Public. They are based on observations from all rounds and show that cooperation increases in Public relative to Personal—i.e., subjects take at later nodes. Moreover,

[27] If pro-self subjects start with pessimistic beliefs about the opponent's behavior at late nodes (Gächter et al., 2008, Sapienza et al., 2013, Butler et al., 2015), they have a further reason to take early, besides their intrinsic motivations.

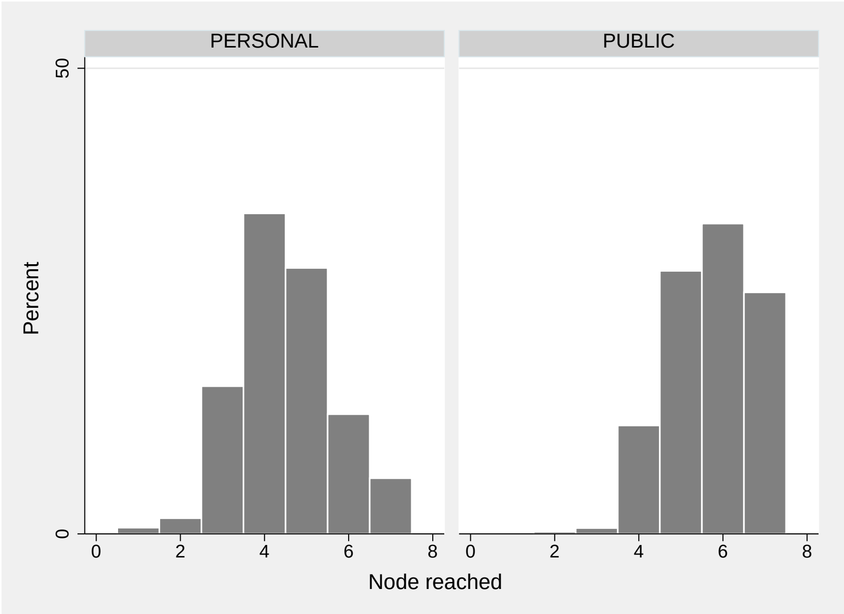

Figure 2: Histograms of the distribution of reached nodes in Personal and Public.

this evidence suggests that the distribution of outcomes is more scattered in Personal, while it is more concentrated in Public. It seems that receiving some aggregate information regarding opponents' play reduces the heterogeneity of behavior.

We proceed with estimating a random-effect panel regression in order to analyze aggregate behavior over time. In

$$z_{it} = \alpha + \beta \cdot t + \gamma \cdot c_{it} + v_i + \epsilon_{it} \tag{1}$$

the terminal node reached by subject i is z, round is t, a vector of control variables is represented by c, the subject-specific error term is v_i and the idiosyncratic one is ϵ_{it}. Only Whites are considered in the regression to avoid double counting. Besides the round number, the regressions include dummies for the experimental conditions. The variable Public equals 1 in the Public condition and Strategy equals 1 in the Strategy condition.

Table 1 reports results for rounds 1-30 (column I), and for rounds 31-40 (column II). During the first 30 rounds of the game the Public condition dummy is positively correlated with the terminal node reached (significant at the 1% level). There is a negative time trend of the terminal node reached in Personal but not in Public. In the last 10 rounds there are no more time trends, while the positive level effect of the Public condition persists.

Hence, in line with previous CG evidence, we find some unraveling when learning possibilities are limited to own experience. While in Personal the average terminal node decreases during rounds 1-30, aggregate behavior stabilizes during the last 10

Table 1: Aggregate behavior over time

	Terminal node			
	I: Rounds 1-30		II: Rounds 31-40	
Public	0.773***	(0.131)	2.460***	(0.680)
Round	-0.0169***	(0.00317)	-0.0136	(0.0162)
Public × Round	0.0301***	(0.00366)	-0.0239	(0.0187)
Strategy	0.0462	(0.131)	0.484	(0.680)
Strategy × Public	-0.202	(0.167)	-0.406*	(0.213)
Strategy × Round	-0.00190	(0.00366)	-0.0150	(0.0187)
Constant	4.848***	(0.0970)	4.765***	(0.584)
Observations	3,840		1,280	

Notes: the dependent variable is the terminal node reached in a match in the first 30 rounds (column I) and the last 10 rounds (column II) respectively; robust standard errors are in parentheses; significance levels are: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

rounds (mean: 4.26). When subjects are informed about the opponent's average behavior (Public) the mean terminal node is substantially higher, it increases during rounds 1-30, and stabilizes during the last 10 rounds (mean: 5.67). Furthermore, the dummy for the Strategy condition is not significantly different from zero. At the aggregate level, subjects do not seem to behave differently as compared to Direct. This is in line with previous evidence that do not find different behavioral patterns across the two methods (Nagel and Tang, 1998, Kawagoe and Takizawa, 2012, Garcia-Pola et al., 2016).

4.2 Effects of aggregate information

We proceed with a more detailed investigation of choices in the last ten rounds (i.e., when behavior stabilized). First of all, we can confirm that our auxiliary assumption (average passing rates at all nodes in all treatments in the last ten rounds are higher than $\frac{1}{3}$) is satisfied. Figure 3 shows the average frequency of choosing Pass for each of the six decision nodes of the game (w_1, b_1, w_2, b_2, w_3, b_3). At each node values are provided for the treatments (Personal-Direct, Personal-Strategy, Public-Direct, Public-Strategy) and for the preference types (pro-self and pro-social).

Three behavioral patterns seem to be present. Virtually all subjects decide to Pass at the first two decision nodes (w_1, b_1), irrespectively of treatment or preference type. At middle nodes (w_2, b_2 and w_3)—and mostly in Strategy—pro-socials tend to Pass more often than pro-selfs in Personal, while in Public both preference types chose Pass at similar rates. At the last node (b_3), pro-socials tend to Pass more often. We will test for the significance of the behavioral differences in a regression framework, taking into account the repeated nature of the observations over rounds as well as the fact that decisions at later nodes depend on those at earlier nodes.

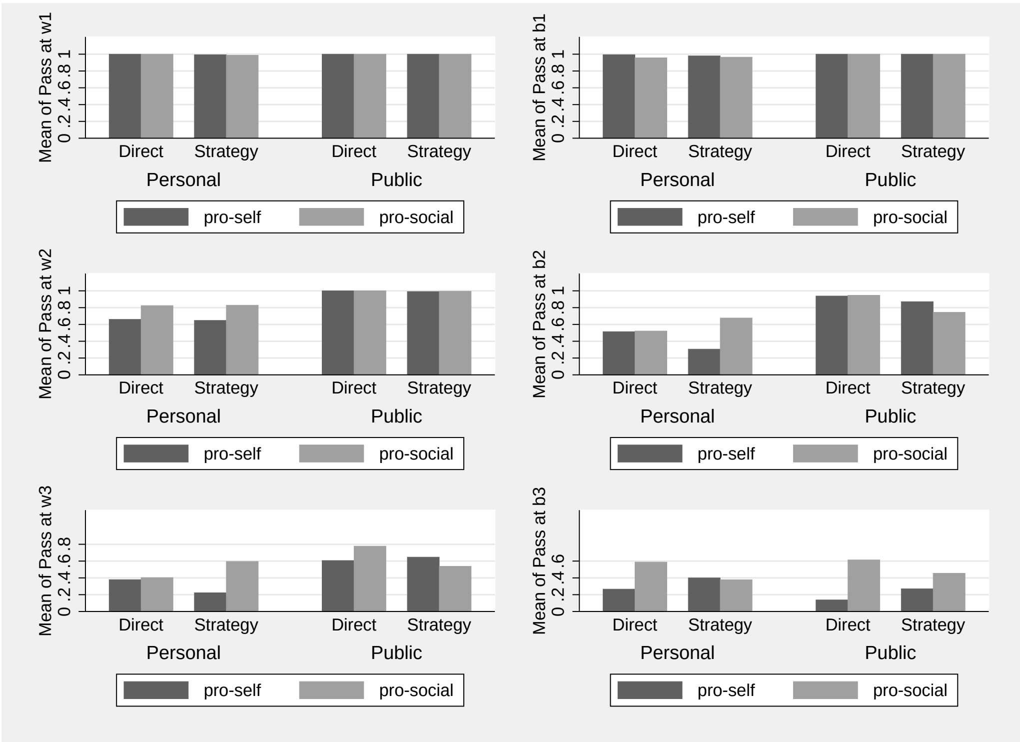

Figure 3: Each of the 6 graphs represents the average frequency of Pass (on the vertical axis) at each decision node of the game—where w_1 (b_1), w_2 (b_2) and w_3 (b_3) are respectively the first, second and third decision node of White (Black)—in the last 10 rounds by information treatment (Personal *vs* Public), by method of play (Direct *vs* Strategy) and by preference type (pro-self *vs* pro-social).

In order to empirically investigate these decision patterns, we focus on the sequence of decisions within a period and set up an econometric model for sequential choices over nodes at that round. To make Strategy data about choices comparable with Direct data, we map selected stopping nodes of Strategy into a sequence of binary decisions at each node.[28] We take into account that decisions at that node can be made only by subjects that have chosen Pass at the preceding nodes, inducing sequential partial observability of choices over nodes; also, we allow for cross-node dependency of choices. As our dependent variable is whether a subject decided to Take (0) or Pass (1), we adopt probit-type specifications for the choice equation at each node and allow for cross-equation correlation of the unobservable determinants of choices, yielding a system of seemingly unrelated regressions. This system of equations approach effectively accounts for choice dependency within a round; we estimate the model pooling data across rounds and correcting standard errors for the presence of repeated observations on the same subject over rounds. One advantage of the approach based on this system of equations is that it enables us testing the joint significance of the treatments at all nodes. Our econometric model follows. In

$$l^*_{in} = \alpha_n + \beta_n \cdot \theta_i + \delta_n \cdot T_i + \epsilon_{in}; \;\; l_{in} = I(l^*_{in} > 0) \tag{2}$$

with

$$n = 1, 2, 3; \;\; \epsilon_i \sim MVN(0, \Sigma)$$

we assume that choices of subject i at each node n are driven by a latent propensity to Pass, denoted l^*, which depends linearly on the individual type θ_i, a vector of treatment dummies T and an error that is assumed to be distributed as a Standard Normal variate. We do not observe the latent variable l^*, but rather its discrete realization l which takes on value 0 (Take) or 1 (Pass) depending on the latent propensity crossing some threshold level that is normalized to 0 without loss of generality. We admit cross-equation correlation by allowing the vector of error terms ϵ_i (consisting of the three errors of each nodes equation) to be distributed as multi-variate normal with correlation matrix Σ. The estimates of cross diagonal elements of this matrix (denoted ρ_{rs}; $r = 1, 2, 3$; $s = 2, 3$; $s > r$) provide a test of cross-node dependency in the unobservable determinants of the choices.[29] As our

[28] Data from Strategy could also be analyzed by means of a regression model for the choice of the stopping node, without the need of seeing such choice as the outcome of a sequence of decisions made at each node. Such an approach leads to stronger results than the ones reported, see Appendix E.

[29] As discussed, nodes 2 and 3 feature sequential partial observability in the sense that the set of subjects for whom we observe choices at these nodes consists only of subjects deciding to Pass at the preceding node (obviously there is no truncation at node 1). Our estimator allows for this feature of the data generating process.

analysis focuses on the last ten rounds we consider a restricted version of the model with only two equations, because in these rounds virtually all subjects chose Pass at their first node, making estimation unfeasible at that node due to lack of variation in the dependent variable.

Table 2 presents regression results, separated by nodes.[30] Coefficients for Public are positive and highly significant at all nodes except at the last one. Coefficients for θ are not significantly different from zero, while the interaction term between θ and Strategy is positive and significant at node b_2. Tests of the joint effect of θ and its interaction with Strategy indicate significance not only at b_2 (Wald test, $p = 0.01$) but also at w_3 ($p = 0.01$). The interaction between θ, Strategy and Public is negative and significant at nodes b_2 (1% level) and w_3 (5% level). Further Wald tests considering all of Whites' nodes confirm the joint significance of θ in Personal-Strategy ($p = 0.01$), while θ is not significantly different from zero in Public-Strategy ($p = 0.94$). Corresponding tests for Blacks' nodes miss significance ($p = 0.12$). This is in line with an analysis of ρ, the correlation of the error terms. For Whites it is significant which supports the approach to treat decisions jointly (Wald test, $p < 0.01$). For Blacks the correlation is not significant, an indication that the choice at the last node is orthogonal to the decisions at earlier nodes.

Summarizing behavior, choices of pro-socials and pro-selfs in Personal-Strategy are distinct at nodes b_2 and w_3, while in Personal-Direct they do not differ significantly.

Result 1 *In Personal-Strategy, behavior in the last 10 rounds is heterogeneous across preference types: pro-social subjects tend to Take at later nodes than pro-selfs.*

In Public, subjects generally tend to choose Take at later nodes. This difference is driven by the behavior of pro-selfs as they tend to Take later in Public than in Personal. In contrast, pro-socials do not exhibit a later tendency to Take in Public-Strategy compared to Personal-Strategy. As a consequence, the heterogeneity of behavior across preference types disappears in Public-Strategy.

Since in Personal-Direct there was no heterogeneity of behavior in the first place, we cannot comprehensively explore whether the release of aggregate information has different effects in Direct and Strategy.

Result 2 *In Public, subjects tend to Pass more often in the last 10 rounds (compared to Personal); in Strategy, this effect is driven by the change of behavior of*

[30]In all our following regressions we focus on sign and significance of the reported results. We do not interpret marginal effects, thus avoiding the resulting complications with respect to interaction terms.

Table 2: Treatments comparison

	w_2		b_2	
θ	0.0090	(0.015)	-0.00061	(0.013)
Strategy	0.082	(0.50)	-0.82	(0.61)
Public	6.64***	(0.37)	1.93***	(0.50)
θ × Public	0.0026	(0.015)	0.0016	(0.018)
θ × Strategy	0.015	(0.020)	0.057**	(0.024)
Strategy × Public	-4.52***	(0.67)	0.70	(0.75)
θ × Strategy × Public	-0.017	(0.028)	-0.076***	(0.028)
Constant	0.076	(0.37)	-0.53	(0.33)
	w_3		b_3	
θ	0.014	(0.014)	0.053*	(0.028)
Strategy	-0.45	(0.56)	1.15	(0.93)
Public	1.20***	(0.41)	-1.16	(0.95)
θ × Public	0.0014	(0.017)	0.0061	(0.023)
θ × Strategy	0.028	(0.021)	-0.070**	(0.033)
Strategy × Public	0.65	(0.67)	0.010	(1.06)
θ × Strategy × Public	-0.053**	(0.026)	0.030	(0.041)
Constant	-1.01***	(0.37)	-0.51	(1.35)
ρ	0.988 ***	0.009	-0.623	0.649
Observations	1,140		1,022	

Notes: System of seemingly unrelated regressions with sequential partial observability; the dependent variables are the choice (Take (0) or Pass (1)) at each decision node of White and Black, in any round t from 31 to 40; in particular, w_2 (b_2), w_3 (b_3), indicate White's (Black's) second and third decision node; robust standard errors in parentheses; significance levels are: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

pro-selfs who take at later nodes.

Given pro-selfs' choices in Personal-Strategy, we now turn to round 40 beliefs data. Do pro-selfs underestimate the true probabilities of passing at the end of the game? For this purpose, we compare the beliefs of pro-selfs and pro-socials and also relate them to actual behavior. Table 3 presents determinants of subjects' last round beliefs. We employ the regression framework introduced earlier but now use the beliefs of subjects as dependent variable. Beliefs are about whether the opponent will play Pass at the next node, that is, the w_1 belief is about the choice at b_1, the b_1 belief is about the choice at w_2, etc. This means beliefs data are not available for b_3. Regressors are θ, the treatment dummies Public and Strategy, and their respective interactions.

In Public, beliefs are higher across the board (except w_1). At w_3, the correlation between θ and beliefs is positive and significant in Strategy, also taking into account the negative main effect of θ (Wald test, $p = 0.03$). The interaction between θ, Strategy and Public is negative and highly significant (but does not reverse the overall effect, Wald test, $p = 0.49$). Across all nodes of White, pro-socials' beliefs in Personal-Strategy are higher than pro-selfs' ($p = 0.06$), while they are not significantly different in Public-Strategy ($p = 0.12$). This pattern is in line with beliefs data from middle rounds (17-19), see Table 8 in Appendix C for regressions results.[31]

Finally, do subjects, especially pro-selfs, estimate actual behavior correctly at the end of the game? In Personal, both preference types substantially underestimate actual behavior starting at w_3 (Strategy, t-tests, $p < 0.05$), respectively at b_2 (Direct, t-tests, $p < 0.05$). In Public instead, neither pro-selfs nor pro-socials underestimate actual choices (at all nodes, both Strategy and Direct). Their beliefs are also not different from each other at any node (t-tests: Strategy, $p > 0.11$; Direct, $p > 0.07$).

We now ask how far subjects are from BNE behavior in each information treatment.[32] Considering the last 10 rounds, we compute the share of subjects in either role who are playing optimally at a certain node given the passing rate of the opponent at the next node.[33] Since at nodes w_1 and b_1 almost all subjects are playing

[31]Already in rounds 17-19 beliefs in Public are consistently higher, except at w_1. Also the correlation between θ and beliefs at w_3 is positive and significant in Strategy. However, it is not considering the negative main effect of θ (Wald test, $p = 0.09$). Similarly to round 40, across all nodes of White pro-socials' beliefs in Personal-Strategy are higher than pro-selfs' ($p = 0.05$), while they are not significantly different in Public-Strategy ($p = 0.94$).

[32]More precisely, we ask how far subjects are from playing at the aggregate level in a way that is observationally equivalent to a BNE.

[33]Notice that we consider what is optimal for a subject and not what is *rational*. That is, we analyze whether subjects are choosing actions with the highest expected payoff given the actual frequencies of the opponent's choices, and not given their subjective probabilities of the opponent's choices (that we have elicited only in round 40).

Table 3: Round 40 beliefs

	w_1		b_1	
θ	-0.00096	(0.0018)	-0.0020	(0.0026)
Public	0.052	(0.037)	0.26***	(0.049)
Strategy	0.0022	(0.040)	0.020	(0.046)
θ × Public	0.0013	(0.0015)	0.0023	(0.0026)
θ × Strategy	0.0023	(0.0015)	0.0026	(0.0027)
θ × Strategy × Public	-0.0024**	(0.0012)	-0.0039*	(0.0023)
Constant	0.93***	(0.053)	0.72***	(0.051)
	w_2		b_2	
θ	-0.0037	(0.0036)	-0.00020	(0.0024)
Public	0.38***	(0.072)	0.43***	(0.067)
Strategy	-0.037	(0.075)	-0.025	(0.066)
θ × Public	0.0052	(0.0034)	-0.00023	(0.0029)
θ × Strategy	0.0054	(0.0038)	0.0012	(0.0028)
θ × Strategy × Public	-0.0085***	(0.0027)	-0.0028	(0.0023)
Constant	0.50***	(0.078)	0.24***	(0.059)
	w_3			
θ	-0.00022	(0.0019)		
Public	0.32***	(0.046)		
Strategy	-0.024	(0.049)		
θ × Public	-0.0010	(0.0019)		
θ × Strategy	0.0049**	(0.0024)		
θ × Strategy × Public	-0.0061***	(0.0021)		
Constant	0.095*	(0.050)		
ρ_{12}	.016	(0.044)	0.315 ***	(0.075)
ρ_{13}	0.369 ***	(0.065)	–	–
ρ_{23}	0.18	(0.098)	–	–
Observations	128		128	

Notes: System of seemingly unrelated regressions with sequential partial observability; the dependent variables are the last round beliefs at each decision node of White (Black) about the choice of Black (White) at the next node; robust standard errors in parentheses; significance levels are: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Pass in the last 10 rounds and at node w_2 we do not find a significant difference in behavior across preference types, we focus on behavior at nodes b_2 and w_3. Moreover, we focus on data from Strategy as this is there we do find behavioral heterogeneity.

Given the payoff structure, for an own payoff maximizing White it is optimal to play Pass at w_3 when the average frequency of Pass at b_3 is larger than $\frac{1}{3}$. Similarly, for an own payoff maximizing Black it is optimal to play Pass at b_2—assuming that the strategy he is playing prescribes Take at the last node—when the average frequency of Pass at w_3 is larger than $\frac{1}{3}$. If it is optimal for a own payoff maximizer to pass at these nodes, then it certainly is for all types.[34]

In the last 10 rounds of Personal, the average passing rate at b_3 satisfies this condition (38.4%); the share of pro-selfs actually playing Pass at w_3 is only 22%, while that of pro-socials is 54%. In the final 10 rounds of Public, the average passing rate at b_3 is 36.7%; the share of pro-selfs playing Pass at w_3 is 65%, while that of pro-socials is 53.8%. In the final 10 rounds of Personal, the average passing rate at w_3 is larger than $\frac{1}{3}$ (43.2%); the share of pro-selfs actually playing Pass at b_2 is 30.6%, while that of pro-socials is 67.7%. In the last 10 rounds of Public, the average passing rate at b_3 is 57.9%; the share of pro-selfs playing Pass at b_2 is 87.14%, while that of pro-socials is 74.44%.

This evidence suggests that in Personal the majority of pro-selfs is playing sub-optimally in the last rounds with respect to actual passing rates, while in Public this tendency reverts. Moreover, while for pro-socials there is no difference in behavior across Personal and Public, pro-selfs' frequency of Pass varies consistently from Personal to Public.

4.3 Evolution of behavior

In this section we study whether the heterogeneity of behavior across preference types that we observed in the last ten rounds originates from the very first rounds of play. Figure 4 shows the evolution of behavior in both Strategy and Direct over the course of the game. While for Strategy we consider the average strategy played, for Direct we consider the average stopping node resulting in a match.[35] Pro-socials'

[34]We do not estimate subjects' utility functions but only have a proxy (θ) for their social preference type. Thus, we cannot study whether in general subjects are playing optimally or not. We can only tell when they are not playing optimally—using as a benchmark the optimal choice of an own payoff maximizer.

[35]Average behavior in Strategy is represented by the mean value of Choice which is a variable that takes value 1 for s^1, 2 for s^2, 3 for s^3 and 4 for s^4. In Direct, average behavior is the mean value of Choice which corresponds to the stopping node that results in a match which is 1 if play stops at the first node of White or Black, 2 if it stops at their second node, 3 if it stops at their third node and 4 if play goes beyond their third node.

average behavior is in red, pro-selfs' in blue and the green dashed line shows average choices.

The figure illustrates graphically that behavior in Strategy is heterogeneous from the very beginning. In Personal, pro-selfs consistently take earlier than pro-socials. While both preference types, on average, tend to take earlier over time, the behavioral gap remains essentially the same. Towards the end of the game behavior of the respective type seems to remain on the same level. In Public, pro-selfs' average behavior starts at about the same level as in Personal. However, after about ten rounds they seem to have adjusted and their average behavior is essentially in line with pro-socials'. Figure 4 also includes a representation of subjects' average beliefs (indicated by crosses), separated by preference type.[36] Both pro-selfs' and pro-socials' initial beliefs start well below actual behavior (in Personal as well as Public). Moreover, the figure suggests that while beliefs in Public generally become very much in line with actual choices, in Direct only the beliefs of pro-socials can be considered accurate at the middle rounds and at round 40. Beliefs of pro-selfs are lower.

In Direct, the behavioral pattern in Personal is quite similar for both preference types. Pro-socials' stopping nodes appear to be slightly higher in the beginning but they quickly converge to the same trajectory as pro-selfs'. In Public, pro-selfs take earlier in the very beginning. Over the course of the game they seem to get closer to the level of pro-socials. The beliefs pattern is similar as in Strategy. Initial beliefs of both preference types are markedly off of early behavior, while they are in line with choices at middle rounds and in the end.

In Strategy, the average strategy in rounds 16 to 20 varies between 3.33 and 3.53 in Public and between 2.87 and 2.95 in Personal. In Direct, the range of the average stopping node is 3.49 and 3.57 in Public and 2.53 and 2.87 in Personal. In none of the treatments there is a significant difference of one round in comparison to the others (t-tests, $p > 0.1$). Thus, the elicitation of beliefs did not seem to have an effect on subsequent behavior.

4.3.1 Initial behavior and beliefs

We proceed with an analysis of the relationship between first round choices, initial beliefs and preferences. The regression set up follows the approach introduced previously. We only look at round 1 data now and add initial beliefs as a further re-

[36] Average beliefs are computed aggregating the conjectured average choice of the opponent for pro-socials and pro-selfs, on the basis of the subjective probabilities of strategies (in Strategy) and the conditional subjective probabilities of actions (in Direct).

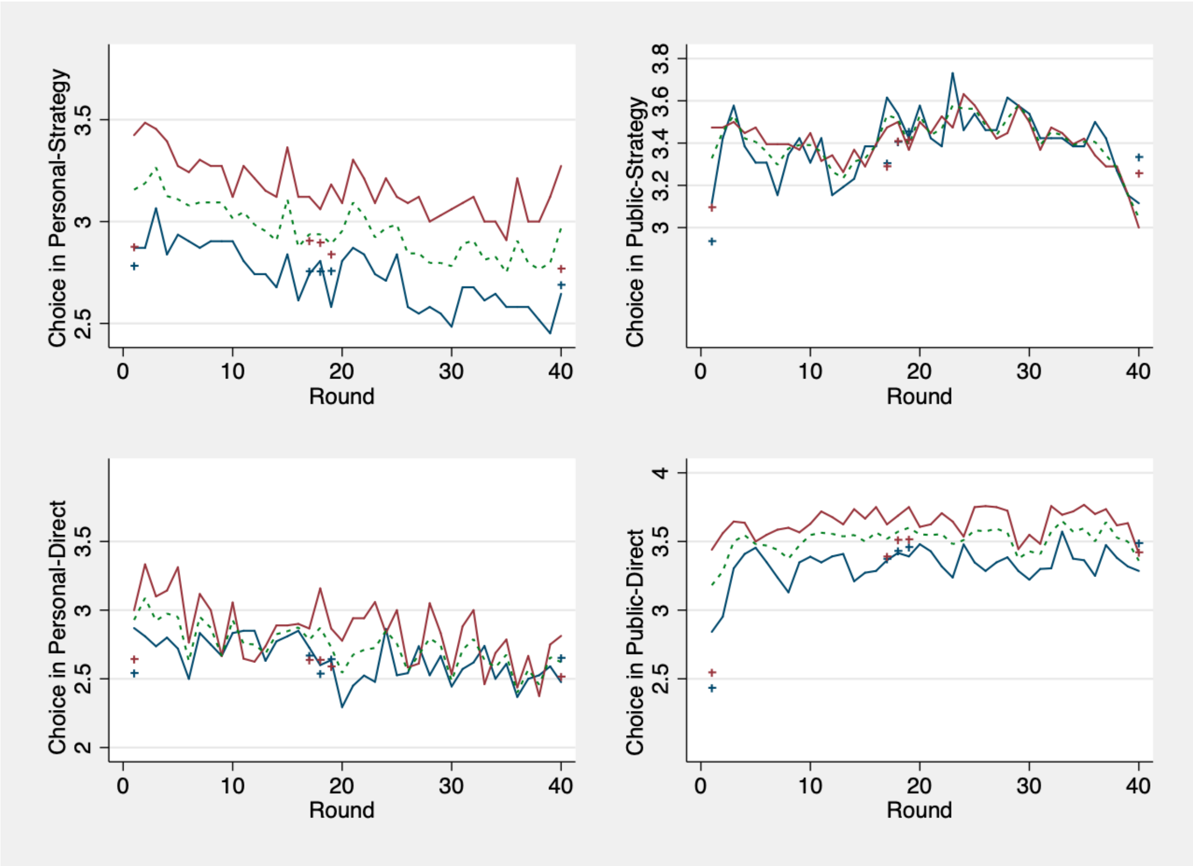

Figure 4: Average behavior of pro-selfs (blue line) and pro-socials (red line) and of all subjects (green line) over the 40 rounds, by method of play (Direct *vs* Strategy) and information treatment (Personal *vs* Public); in Strategy the variable Choice that takes value 1, 2, 3 and 4 when respectively s^1, s^2, s^3 and s^4 are chosen; in Direct Choice takes values 1, 2, 3 when play stops at the first, second and third node of either player, and 4 when they never stop. Crosses represent average beliefs of pro-selfs (in blue) and of pro-socials (in red) that aggregate the conjectured average choice of the opponent of every subject in either group.

gressor. In particular, in each regression we estimate the effect of the belief that the opponent will play Pass at node x on the choice to Pass at node $x-1$. Consequently, beliefs data are not available for the choice at b_3.

Table 4 presents results. Coefficients of θ are significant at b_3 at the 5% level and the interaction between θ and Strategy at b_2 (1% level). Tests across nodes confirm the joint significance of θ in Direct (Wald tests, $p < 0.05$) as well as in Strategy ($p < 0.05$). Beliefs are positively correlated with choices at nodes w_2 and b_2 (1% level). At b_2, the interaction term between θ and beliefs is negative and significant at the 5% level. Error term correlations are significant for Whites, thus, supporting the system of equations approach.

Thus, in first round behavior we find initial heterogeneity across our social preferences measure θ. Furthermore, the round 1 analysis indicates that besides preferences also beliefs are a determinant of behavior. The negative interaction suggests that this effect of beliefs is less pronounced for pro-socials than for pro-selfs. When subjects decide whether to Pass, high beliefs can compensate for the lack of social preferences.

Finally, for what regards initial heterogeneity of beliefs, our data suggest a correlation between initial beliefs and preferences, in line with false consensus (Ross et al., 1977). We employ the familiar approach of estimating a system of seemingly unrelated regressions with sequential partial observability.[37] Results (see Table 5) confirm that pro-socials tend to have higher initial beliefs than pro-selfs at node b_2 (5% level). Wald tests considering all nodes confirm joint significance of θ for Blacks ($p = 0.05$) but not for Whites ($p = 0.12$).

Our overall beliefs data (round 1, 17-19 and 40) indicate two different patterns. In round 1, Blacks' beliefs are heterogeneous across preference types but not Whites'. In round 40, we find type-dependence of Whites' beliefs in Personal-Strategy, most prominent at w_3. This last round heterogeneity of beliefs across types can already be observed in the middle rounds, while there is no difference of beliefs across preference types among Blacks. It seems that pro-selfs' learning after about the half way point is scarce and does not result in closing the gap to pro-socials.

Thus, node-specific initial heterogeneity—compatible with a false consensus—seems to fade away and the intuition that biased initial beliefs may exacerbate the type-dependence of the learning process does not find support in our data. Instead, as

[37] We do not include interactions with θ like in beliefs estimations during the game, because we expect the manipulation of information and the method of play to have no differential effect on ex ante beliefs of preference types. We find that round 1 beliefs are not correlated with the Public condition but with Strategy. However, this apparent effect of the different elicitation procedure disappears in later rounds.

Table 4: Individual behavior and preferences in round 1

	w_1		b_1	
θ	0.011	(0.0070)	0.011*	(0.0064)
Initial belief	0.52*	(0.28)	0.30	(0.25)
θ × Initial belief	-0.013*	(0.0076)	-0.011	(0.0087)
Public	0.019	(0.013)	-0.0047	(0.028)
Strategy	0.24	(0.15)	-0.16	(0.21)
θ × Strategy	-0.000087	(0.00064)	0.0033	(0.0047)
Initial belief × Strategy	-0.27	(0.17)	0.086	(0.18)
Constant	0.52**	(0.26)	0.70***	(0.18)
	w_2		b_2	
θ	0.010*	(0.0060)	-0.0014	(0.0098)
Initial belief	0.64***	(0.20)	0.86***	(0.28)
θ × Initial belief	-0.012*	(0.0071)	-0.017**	(0.0078)
Public	-0.0016	(0.044)	0.042	(0.067)
Strategy	-0.43**	(0.19)	-0.43**	(0.18)
θ × Strategy	0.0018	(0.0037)	0.015***	(0.0056)
Initial belief × Strategy	0.40*	(0.21)	0.035	(0.21)
Constant	0.49***	(0.18)	0.68**	(0.29)
	w_3		b_3	
θ	-0.0013	(0.0078)	0.018**	(0.0074)
Initial belief	-0.22	(0.51)		
θ × Initial belief	0.0029	(0.018)		
Public	-0.072	(0.11)	0.032	(0.12)
Strategy	0.20	(0.28)	0.35	(0.23)
θ × Strategy	-0.0019	(0.0098)	-0.016	(0.010)
Initial belief × Strategy	0.073	(0.42)		
Constant	0.69***	(0.24)	-0.015	(0.15)
ρ_{12}	0.23	(0.20)	-0.23	(0.66)
ρ_{13}	0.025	(0.19)	-0.038	(0.33)
ρ_{23}	-.58 ***	(0.21)	-0.325	(0.29)
Observations	128		126	

Notes: System of seemingly unrelated regressions with sequential partial observability; the dependent variables are the choice (Take (0) or Pass (1)) at each decision node of White and Black in round 1; in particular, w_1 (b_1), w_2 (b_2), indicate White's (Black's) first and second decision node; robust standard errors in parentheses; significance levels are: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

the learning process unfolds, a discrepancy between pro-socials' and pro-selfs' beliefs at w_3 (i.e., about behavior at the last node of Black) arises. Not surprisingly this occurs in the treatment where we do observe heterogeneity of behavior in the middle nodes of the game (Personal-Strategy).

Table 5: Round 1 beliefs

	w_1		b_1	
θ	0.0026*	(0.0013)	0.0016	(0.0015)
Public	0.041	(0.035)	-0.070*	(0.040)
Strategy	0.080**	(0.035)	0.21***	(0.040)
Constant	0.75***	(0.058)	0.62***	(0.051)
	w_2		b_2	
θ	0.0028	(0.0018)	0.0040**	(0.0019)
Public	0.11**	(0.044)	0.062	(0.055)
Strategy	0.18***	(0.045)	0.12**	(0.054)
Constant	0.42***	(0.066)	0.22***	(0.055)
	w_3			
θ	0.00023	(0.0017)		
Public	-0.049	(0.045)		
Strategy	0.14***	(0.046)		
Constant	0.20***	(0.061)		
ρ_{12}	.429 ***	(0.083)	0.445 ***	(0.075)
ρ_{13}	0.035	(0.089)	–	–
ρ_{23}	0.418 ***	(0.077)	–	–
Observations	128		128	

Notes: System of seemingly unrelated regressions with sequential partial observability; the dependent variables are the beliefs in the first round at each decision node of White (Black) about the choice of Black (White) at the next node; robust standard errors in parentheses; significance levels are: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

4.3.2 Information and round by round behavior

In order to better understand the effect of aggregate information release on different preference types' behavior, we now consider the effect of information (gained through own experience as well as public disclosure of aggregate choices) on behavior in our analysis.

Behavior in any round $t > 1$ may depend on observations made up to t about behavior of the opponents. The variable *Observed Behavior* represents the personal statistics of a subject at any time t. It consists of the average frequency of Pass at the next node out of all own observations made up to t of behavior of subjects in the opponent's role. If at t a subject has never observed the opponent's choice at the next node, then Observed Behavior does not exist. If a subject observed Pass as often as Take, then Observed Behavior equals 0.5. The presentation of results follows the same fashion as introduced earlier, that is, we estimate a system of equations, with the choice to Pass as the dependent variable. The familiar regressors are amended with *Observed Behavior*. Due to a lack of variation in the data at w_1 this node is skipped.

Table 6: Individual behavior and information in Personal

	b_1	w_2	b_2	w_3
Observed Behavior	5.037*	6.621***	5.220***	8.096***
	(2.743)	(2.283)	(1.184)	(2.952)
θ	0.0267	0.0247	0.0432***	0.0121
	(0.0754)	(0.0616)	(0.0157)	(0.0147)
Observed Behavior $\times$ θ	-0.0565	-0.0115	-0.117***	-0.208**
	(0.0860)	(0.0894)	(0.0417)	(0.106)
Strategy	-7.900**	2.233	-1.240	-0.456
	(3.154)	(2.102)	(1.063)	(0.515)
θ $\times$ Strategy	0.335***	0.0584	0.0344	0.0171
	(0.117)	(0.0786)	(0.0408)	(0.0224)
Observed Behavior $\times$ Strategy	8.636**	-3.246	0.0251	-5.730*
	(3.492)	(3.464)	(2.278)	(3.441)
θ $\times$ Observed Behavior $\times$ Strategy	-0.360***	-0.0786	0.0328	0.180
	(0.126)	(0.124)	(0.0890)	(0.123)
Constant	-1.530	-3.895***	-1.587***	-0.815**
	(2.473)	(1.484)	(0.381)	(0.408)
ρ_{23}			0.996	0.059
			(0.011)	(0.39)
Observations	1,761	2,133	1,761	2,133

Notes: System of seemingly unrelated regressions with sequential partial observability; the dependent variables are the choice (Take (0) or Pass (1)) at each decision node of White and Black, in any round t from 2 to 40; robust standard errors in parentheses; significance levels are: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 6 reports regression results (error term correlation significant for Blacks, Wald test, $p < 0.05$). The main effect of Observed Behavior is positive and significant at the 1% level (except at b_1), while θ is positive and significant (1% level) at b_2. Moreover, the interaction term between θ and Observed Behavior is negative and significant at b_2 (1% level) and w_3 (5% level). Hence, pro-socials exhibit a tendency to Pass later, irrespectively of the behavior they observe, while pro-selfs seem to condition their choice on Observed Behavior. We observe this behavioral pattern also at b_1 but only in Strategy.

In Public, besides the personal database generated by own observations, there is another source of information. After playing, subjects receive aggregate statistics about how subjects in the opponent's role played the game in the past round. This variable, *Aggregate Information*, has a different informational content in the two methods of play. In Direct, it is the average frequency of Pass at any node out of *all observations* actually made overall, in all matches, up to t at that node. In Strategy, it corresponds to the average frequency of Pass at any node out of *all choices* made overall, in all matches, up to t at that node.[38] Notice that in Strategy a choice made at a certain node by a subject in a given match does not need to coincide with the observation made by the opponent in that match as that node may be off the actual path of play.

As the system of equations approach did not result in any error term correlations across the nodes, we turn to a set of single probit regressions, one for each node. The dependent variable is whether a subject decided to Take (0) or Pass (1) at a respective node. Due to the fact that behavior at the first node is always Pass and that Aggregate Information is not available at b_3 these two nodes are not considered.

In Direct, there is no correlation between Aggregate Information and the tendency to Pass, while the interaction term between Aggregate Information and Strategy is significant at all nodes. Taking the negative main effects into account Wald tests confirm a significantly positive effect at b_2 and w_3 ($p < 0.05$) but no significance at earlier nodes ($p > 0.405$). At nodes w_2, b_2 and w_3, Observed behavior is positive and significant at the 1% level, while θ is at b_2 (5% level) and w_3 (1% level). As an alternative regression specification we included interaction terms between *Aggregate Information* and θ/*Observed Behavior*. However, these interaction terms were not significant.

To summarize, we find that subjects generally internalize observed behavior and react to it, irrespective of the treatment. When subjects are informed about strategies (not only actions), the effect of *Aggregate Information* appears to matter. Moreover, we find that pro-selfs have a general tendency to respond to observed behavior, while pro-socials tend to Pass irrespective of what they observe.

[38]In Strategy the exact message that subjects receive consists of the average frequencies of strategies (i.e., stopping nodes) played, on average, by subjects in the opponent's role. However, for the data analysis, we decomposed average frequencies of strategies into average frequencies of planned actions at nodes compatible with the strategies chosen.

Table 7: Individual behavior and information in Public

	b_1	w_2	b_2	w_3
Aggregate Information	-8.81	-7.45	-0.29	-0.41
	(5.61)	(4.70)	(1.85)	(2.07)
Observed Behavior	–	2.82***	3.30***	2.83***
	–	(0.74)	(0.42)	(0.47)
θ	0.024*	0.019*	0.026**	0.027***
	(0.013)	(0.011)	(0.010)	(0.010)
Strategy	8.39*	6.10	-1.91	1.51
	(4.43)	(4.20)	(1.47)	(1.12)
Aggregate Information × Strategy	5.72***	3.94***	4.73*	3.50**
	(0.55)	(1.12)	(2.42)	(1.65)
θ × Strategy	0.0077	-0.020	-0.037**	-0.033
	(0.024)	(0.015)	(0.018)	(0.021)
Observed Behavior × Strategy	–	0.063	-0.79	-1.00
	–	(0.93)	(0.73)	(0.64)
Constant	-1.53	-2.94**	-0.57	-2.64***
	(1.67)	(1.16)	(1.15)	(0.80)
Observations	2,062	2,492	2,484	2,262

Notes: Set of probit regressions; the dependent variables are the choice (Take (0) or Pass (1)) at each decision node of White and Black, in any round t from 2 to 40; at b_1 433 observations could not be used to avoid a lack of variation; robust standard errors in parentheses; significance levels are: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

4.4 Discussion

In the following we discuss the robustness of our results, their relation to existing literature and potential determinants for the lack of heterogeneity in Personal-Direct.

Our regression results use the continuous variable θ as the measure of a subject's pro-sociality. Replacing θ with the binary preference type variable in the regressions by and large does not change results qualitatively (sign and significance levels remain the same), except the regressions presented in table 4 (round 1 choices) and 5 (behavior in Personal). In round 1 the interaction between the preference type and beliefs is not significant ($p = 0.23$) anymore and the preference type is not jointly significant anymore for Blacks ($p < 0.23$). Observed behavior in Personal at b_1 is now also significant at the 1% level in Direct, while the interaction between observed behavior and the preference type at w_3 is not significant ($p = 0.12$) anymore. Neither of these differences are central to our tested hypotheses. As an alternative to the SVO-based preference types we developed a categorization that derives from choices in the Trust game (see Appendix B for details). The only qualitative difference in regression results is that round 1 beliefs have no significant effect on initial behavior ($p = 0.13$).

In our experiment, aggregate behavior is similar across the two methods of play, in line with related studies (Nagel and Tang, 1998, Kawagoe and Takizawa, 2012, Garcia-Pola et al., 2016). When we provide ex post information, a substantial positive effect on average payoffs results. In particular, public information release increases passing rates at middle nodes, while it does not affect much cooperation at the very last node. As feedback informs subjects that passing rates at the last nodes are higher than they expected, beliefs are revised upwards and behavior adapts. The direction of the effect of aggregate information on cooperation is not the focus of our paper, though.[39] We are indeed more interested in the mechanisms through which different ex post information structures affect behavior of different preference types than on the sign of the effect *per se*.

In the following we look in more detail at the area where behavior in our experiment deviates from what we hypothesized. At first glance it seems puzzling that behavior at middle nodes differs across preference types in Personal-Strategy, while in Personal-Direct preference types behave quite similarly. What could be an explanation for the lack of heterogeneity in Personal-Direct?

[39] Maniadis (2012) shows that the effect of a public feedback on outcomes varies significantly with the payoff structure. By increasing only the payoff from Pass of the player active at the very last node, the effect of the public feedback on average payoffs, negative in the benchmark (more unraveling), becomes positive (more cooperation).

When the game structure is essentially static and subjects are asked to express a choice over terminal nodes (Strategy), distributional concerns with respect to payoffs are a natural determinant of behavior (in addition to expectations). When the game structure is explicitly interactive and dynamic (Direct), besides distributional concerns, other intrinsic motives, induced by emotions, may kick in.[40] Emotions may be triggered by the fact that in Direct subjects observe how their opponent is *actually* playing. Hence, subjects who do not have any particular concern for the payoff of a generic other may still be inclined to reciprocate the opponent for choosing Pass. Similarly, guilt averse subjects may infer that behind the opponent's choice of passing there is the expectation of being rewarded and so they feel obliged to also pass when it is their turn to move. Thus, there may be subjects classified as pro-selfs, who behave similarly to pro-socials in Direct.

While these reactions would have a positive effect on the tendency to Pass, the dynamic interaction in Direct could also have negative emotional effects. Expectations about the opponent's Pass choice may build up by the exchange of Pass choices at the early nodes. The disappointment by an opponent choosing Take (right after one decided to Pass once more) would consequently be higher in Direct than in Strategy. Such unfulfilled expectations would result in negative emotions like frustration and anger (see Battigalli et al., 2019).[41] Stronger negative emotions may trigger more aggressive behavior in the next round.[42] More intense reactions to negative emotions is what could explain why pro-socials tend to pass less in Direct than in Strategy.

Summing up, as our preferences measures are based on *distributional* concerns they do not capture the full range of potential social preferences. Hence, the lack of heterogeneity in Personal-Direct might be due to subjects' emotional reactions in Direct.

5 Conclusion

In this paper we studied experimentally whether persistent behavioral heterogeneity in the Centipede game (CG) can be explained by heterogeneous social preferences

[40]According to Brandts and Charness (2011), the direct response method might trigger emotional reactions that are absent in the strategy method. Notice however that in our experiment the direct method and the variant of the strategy method that we adopt do not produce different results at the aggregate level. What changes across methods is the impact of social preferences on behavior.

[41]Notice according to the definition of "simple guilt"by Battigalli et al. (2019) retaliation as an expression of frustration may occur towards a generic other and not necessarily the one who caused such frustration, in a similar vein to indirect reciprocity.

[42]See Harth and Regner (2017) for related evidence of anger spillovers to behavior in following rounds in a Trust game setting.

and type-dependent conjectures about the opponent's play. Our experimental analysis shows that last rounds behavioral heterogeneity can depend on the type of information feedback about the opponent's behavior. When feedback is limited to own outcomes, a stable configuration where some types (pro-selfs) take earlier than others (pro-socials) can emerge. When feedback is independent of own play and is about the aggregate behavior of subjects in the opponent's role, heterogeneity of behavior and beliefs drops down, with pro-selfs behaving more similarly to pro-socials.

A first important contribution of this paper is that the type of information feedback can alter the long run distribution of strategies in the CG. This result is based on the intuition—confirmed by our data—that the extent of uncertainty about the opponent's behavior can be endogenous to own play and thus to own preferences. Hence, an exogenous feedback that resolves agents' uncertainty about the opponent's average behavior can alter the equilibrium play of some of them. This innovates on previous CG studies which considered only the direct effect of preferences on behavior in incomplete information settings (see, e.g., McKelvey and Palfrey, 1992, Healy, 2016). We conclude that preferences-dependent conjectures can be an additional source of persistent heterogeneous behavior. Moreover, by showing that heterogeneity across preference types in the CG can emerge as an equilibrium phenomenon due to individual learning, we also complement the literature (e.g. Fey et al., 1996, Garcia-Pola et al., 2016) that focuses on initial heterogeneity and finds mixed evidence about the role of social preferences in this game.

Our analysis also delivers a methodological contribution. We show that the characteristics of stable outcomes resulting from recurrent interactions crucially depend on the ex post information structure. Our results illustrate that depending on the information conditions long run behavior can be better described by solution concepts weaker than Bayesian Nash equilibrium. Concepts that admit heterogeneous conjectures, like self-confirming equilibrium (Battigalli, 1987, Fudenberg and Levine, 1993a), are more appropriate when the feedback about the opponent's behavior is not informative enough to make all agents learn the opponent's strategies. Importantly, what experimenters can expect from long run outcomes of recurrent interactions depends on the type of feedback that they provide to subjects. For example, we show that heterogeneous preferences can rationalize long run outcomes when the feedback is not informative enough, but they cannot when subjects receive detailed aggregate statistics, which are exogenous with respect to their own play.

Finally, our results have general implications that go beyond understanding behavior in the CG. They apply to other dynamic social dilemmas where social preferences and sequential rationality matter. A natural application are dynamic games involving trusting decisions, but it can also be applied to all dynamic games where what a player can observe ex post about the opponent's behavior depends on his own strategy. The preference types that might be relevant obviously depend on the strategic context. In the CG we picked distributional concerns as a straightforward way to classify subjects. We discussed the possible reasons why such preferences may not explain behavior in sequential settings. Yet, this does not undermine the central message: when heterogeneous preferences result in heterogeneous behavior, whether we observe persistent behavioral heterogeneity crucially depends on the information feedback. Hence, what is essential is defining what players can observe ex post and whether the feedback is informative enough to make them learn the opponent's strategies independently of how they are playing.

Acknowledgments

We are grateful to the Max Planck Institute of Economics, Jena, for financial support. We thank two anonymous referees, Pierpaolo Battigalli, Lorenzo Cappellari, Paolo Crosetto, Anna De Paoli, Werner Güth, Zacharias Maniadis, Natalia Montinari, Ariel Rubinstein, Luca Stanca and attendants of the ESA European Conference 2012 (Cologne), of UECE 2012 (Lisbon), of the *Game Theory at the Universities of Milano III* 2013 workshop (Milano), and seminar participants at IMW 2013 (Bielefeld) for useful suggestions. The usual disclaimer applies.

References

Adriani, F. and Sonderegger, S.: 2015, Trust, trustworthiness and the consensus effect: An evolutionary approach, *European Economic Review* **77**, 102–116.

Anderhub, V., Engelmann, D. and Güth, W.: 2002, An experimental study of the repeated trust game with incomplete information, *Journal of Economic Behavior & Organization* **48**(2), 197–216.

Armantier, O. and Treich, N.: 2013, Eliciting beliefs: Proper scoring rules, incentives, stakes and hedging, *European Economic Review* **62**, 17–40.

Azrieli, Y., Chambers, C. P. and Healy, P. J.: 2018, Incentives in experiments: A theoretical analysis, *Journal of Political Economy* **126**(4), 1472–1503.

Battigalli, P.: 1987, Comportamento razionale ed equilibrio nei giochi e nelle situazioni sociali, *unpublished undergrad dissertation, Bocconi University, Milano* .

Battigalli, P., Cerreia-Vioglio, S., Maccheroni, F. and Marinacci, M.: 2015, Self-confirming equilibrium and model uncertainty, *The American Economic Review* **105**(2), 646–677.

Battigalli, P., Dufwenberg, M. and Smith, A.: 2019, Frustration, aggression, and anger in leader-follower games, *Games and Economic Behavior* .

Battigalli, P. and Guaitoli, D.: 1997, *Conjectural equilibria and rationalizability in a game with incomplete information*, Springer.

Berg, J., Dickhaut, J. and McCabe, K.: 1995, Trust, reciprocity, and social history, *Games and economic behavior* **10**(1), 122–142.

Blanco, M., Engelmann, D., Koch, A. K. and Normann, H.-T.: 2010, Belief elicitation in experiments: is there a hedging problem?, *Experimental Economics* **13**(4), 412–438.

Brandts, J. and Charness, G.: 2011, The strategy versus the direct-response method: A first survey of experimental comparisons, *Experimental Economics* **14**(3), 375–398.

Brown, A. L. and Healy, P. J.: 2018, Separated decisions, *European Economic Review* **101**, 20–34.

Brown, G. W.: 1951, Iterative solution of games by fictitious play, *Activity analysis of production and allocation* **13**(1), 374–376.

Butler, J., Giuliano, P. and Guiso, L.: 2015, Trust, values and false consensus, *International Economic Review* **56**, 889–915.

Charness, G., Gneezy, U. and Halladay, B.: 2016, Experimental methods: Pay one or pay all, *Journal of Economic Behavior & Organization* **131**, 141–150.

Cox, J. C. and James, D.: 2012, Clocks and trees: Isomorphic dutch auctions and centipede games, *Econometrica* **80**(2), 883–903.

Danz, D., Huck, S. and Jehiel, P.: 2016, Public statistics and private experience: Varying feedback information in a take-or-pass game, *German Economic Review* **17**(3), 359–377.

Dekel, A., Fudenberg, D. and Levine, D. K.: 2004, Learning to play Bayesian games, *Games and Economic Behavior* **46**, 282–303.

Embrey, M., Fréchette, G. R. and Yuksel, S.: 2017, Cooperation in the finitely repeated prisoners dilemma, *The Quarterly Journal of Economics* **133**(1), 509–551.

Fey, M., McKelvey, R. D. and Palfrey, T. R.: 1996, An experimental study of constant-sum centipede games, *International Journal of Game Theory* **25**(3), 269–287.

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

Fudenberg, D. and Kreps, D. M.: 1995, Learning in extensive-form games i. self-confirming equilibria, *Games and Economic Behavior* **8**(1), 20–55.

Fudenberg, D. and Levine, D. K.: 1993a, Self-confirming equilibrium, *Econometrica* pp. 523–545.

Fudenberg, D. and Levine, D. K.: 1993b, Steady state learning and nash equilibrium, *Econometrica* pp. 547–573.

Gächter, S., Nosenzo, D., Renner, E. and Sefton, M.: 2008, Who makes a good leader? social preferences and leading-by-example.

Gamba, A.: 2013, Learning and evolution of altruistic preferences in the centipede game, *Journal of Economic Behavior & Organization* **85**, 112–117.

Garcia-Pola, B., Iriberri, N. and Kovarik, J.: 2016, Non-equilibrium play in centipede games.

Greiff, M., Ackermann, K. and Murphy, R.: 2018, Playing a game or making a decision? methodological issues in the measurement of distributional preferences, *Games* **9**(4), 80.

Greiner, B.: 2004, The online recruitment system orsee 2.0 - a guide for the organization of experiments in economics, *mimeo*, Department of Economics, University of Cologne.

Harth, N. S. and Regner, T.: 2017, The spiral of distrust:(non-) cooperation in a repeated trust game is predicted by anger and individual differences in negative reciprocity orientation, *International Journal of Psychology* **52**, 18–25.

Healy, P. J.: 2016, Epistemic experiments: Utilities, beliefs, and irrational play.

Jehiel, P.: 2005, Analogy-based expectation equilibrium, *Journal of Economic Theory* **123**(2), 81–104.

Kamada, Y.: 2010, Strongly consistent self-confirming equilibrium, *Econometrica* **78**(2), 823–832.

Kawagoe, T. and Takizawa, H.: 2012, Level-k analysis of experimental centipede games, *Journal of Economic Behavior & Organization* **82**(2-3), 548–566.

Kreps, D. M., Alt, J. E. and Shepsle, K. A.: 1996, Corporate culture and economic theory, *Firms, Organizations and Contracts, Oxford University Press, Oxford* pp. 221–275.

Maniadis, Z.: 2012, *Aggregate information and the centipede game: a theoretical and experimental study*, University of Southampton, School of Social Sciences, Economics Division.

McKelvey, R. D. and Palfrey, T. R.: 1992, An experimental study of the centipede game, *Econometrica* pp. 803–836.

Mentzakis, E. and Mestelman, S.: 2013, Hypothetical bias in value orientations ring games, *Economics Letters* **120**(3), 562–565.

Mermer, A. and Suetens, S.: 2017, Choosing to be informed in a repeated trust game, *Technical report*, mimeo.

Murphy, R. O., Ackermann, K. A. and Handgraaf, M.: 2011, Measuring social value orientation, *Judgment and Decision Making* **6**(8), 771–781.

Nagel, R. and Tang, F. F.: 1998, Experimental results on the centipede game in normal form: an investigation on learning, *Journal of Mathematical psychology* **42**(2-3), 356–384.

Ross, L., Greene, D. and House, P.: 1977, The false consensus effect: An egocentric bias in social perception and attribution processes, *Journal of experimental social psychology* **13**(3), 279–301.

Rutström, E. E. and Wilcox, N. T.: 2009, Stated beliefs versus inferred beliefs: A methodological inquiry and experimental test, *Games and Economic Behavior* **67**(2), 616–632.

Sapienza, P., Toldra-Simats, A. and Zingales, L.: 2013, Understanding trust, *The Economic Journal* **123**(573), 1313–1332.

Schotter, A. and Trevino, I.: 2014, Belief elicitation in the laboratory, *Annu. Rev. Econ.* **6**(1), 103–128.

Selten, R.: 1967, Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments, *Beiträge zur experimentellen Wirtschaftsforschung* **1**, 136–168.

Wang, R. Y. and Ng, C. N.: 2015, Can centralized sanctioning promote trust in social dilemmas? a two-level trust game with incomplete information, *PloS one* **10**(4), e0124513.

Appendices

A. Social Value Orientation (SVO)

The SVO slider measure (Murphy et al., 2011) consists of six primary items and nine optional ones. All items have the same general form. In each item subjects face a resource allocation choice over a well defined continuum of joint payoffs (self and other). One item, for instance, features the trade-off between the perfectly individualistic choice of (100, 50) and the perfectly altruistic choice of (50, 100). Besides these extreme values there are always seven in-between allocations to allow for intermediate choices. The remaining five primary items are (85, 85) vs. (85, 15), (85, 15) vs. (100, 50), (50, 100) vs. (85, 85), (50, 100) vs. (85, 15) and (100, 50) vs. (85, 85). In contrast to the simple categorization of previous SVO measures the slider measure yields a continuous measure based on choices in the six primary items, the SVO angle. It equals $arctan\frac{\pi_s-50}{\pi_o-50}$, where π_s is the average monetary payoff allocated to self and π_o is the average monetary payoff allocated to the other person. See Figure 5 for the distribution among our subjects (mean angle 22.37, standard deviation 13.29). The social value orientation literature distinguishes four idealized social orientations among individuals and Murphy et al. (2011) derive thresholds to separate the four types from each other: competitive (less than -12.04), individualistic (between -12.04 and 22.45), pro-social (between 22.45 and 57.15), and altruistic (more than 57.15). A relatively even share of the two main types (pro-social and individualistic) is common, see Murphy et al. (2011) but also Mentzakis and Mestelman (2013), Greiff et al. (2018).

B. Trust game

The Trust game played in part 1 of the experiment is illustrated in Figure 6. Player A has to decide whether to invest an endowment of 6 experimental currency units (ECU), by sending it to player B (action a_2), or keep the money for himself (action

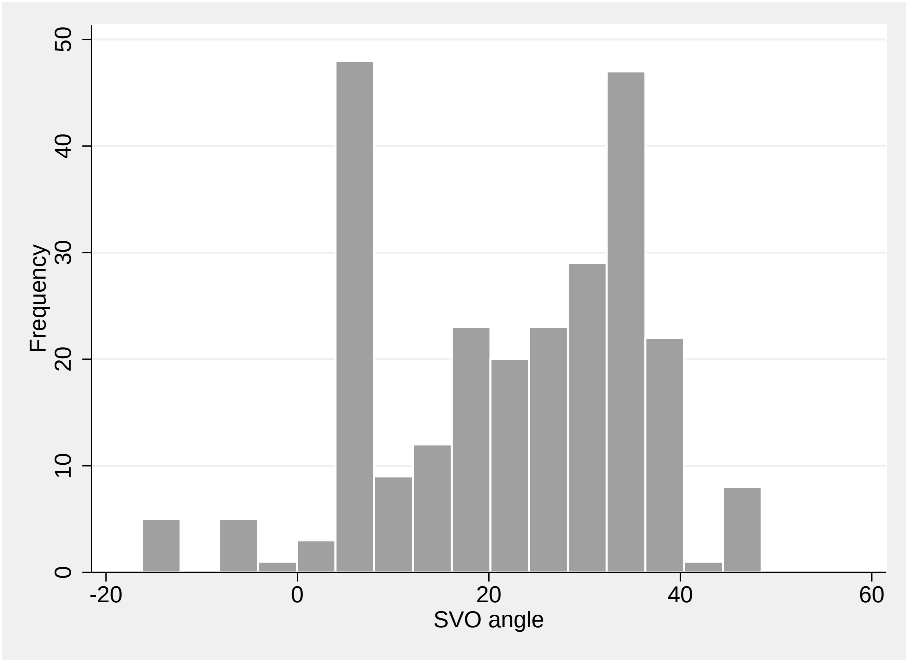

Figure 5: Distribution of Social Value Orientation scores (slider measure)

a_1). If A sends the 6 ECU, the total amount available to the players increases and it is B's turn to move, while if A keeps the 6 ECU, the game ends. In the latter case, A has also to decide on a (small) voluntary transfer t to B, $t \in \{0, 0.5, 1, 1.5, 2\}$, so that A gets $6 - t$ ECU, while B gets t ECU. This transfer represents a sort of compensation to B for stopping the game. If, instead, A sends the money to B, B has to decide whether to further invest the augmented amount (action b_2) or to keep most of it (action b_1). If B chooses b_1, he gets 6 ECU, while A gets 2 ECU. If, instead, B further invests, the total amount available to the players increases, but B only gets 4 ECU out of it, while A gets 8 ECU. The game is played with the strategy method: subjects are first requested to make a decision in B's shoes and then in A's shoes. Subjects who chose a_1 have also to choose the transfer t. They are assigned either role A or B, they are anonymously matched with another subject in the same session and paid the monetary outcomes generated by the strategy profile they adopted (at the end of the experiment).

We chose this particular payoff structure for our Trust game to maintain similarities with the CG payoff structure. In both games whenever a player passes money to the opponent the surplus increases, but this may occur at the expenses of his material payoff and the payoff allocation when B passes (b_2) is more efficient than that after take (b_1).[43] Moreover, like in the CG, B obtains 2 ECUs less when honoring trust—

[43]Notice that in the original version of the Trust game by Berg et al. (1995), the surplus generated by trust is tripled with respect to the initial surplus and, as in many experiments thereafter, the augmented surplus stays constant across the outcomes that follow B's decision.

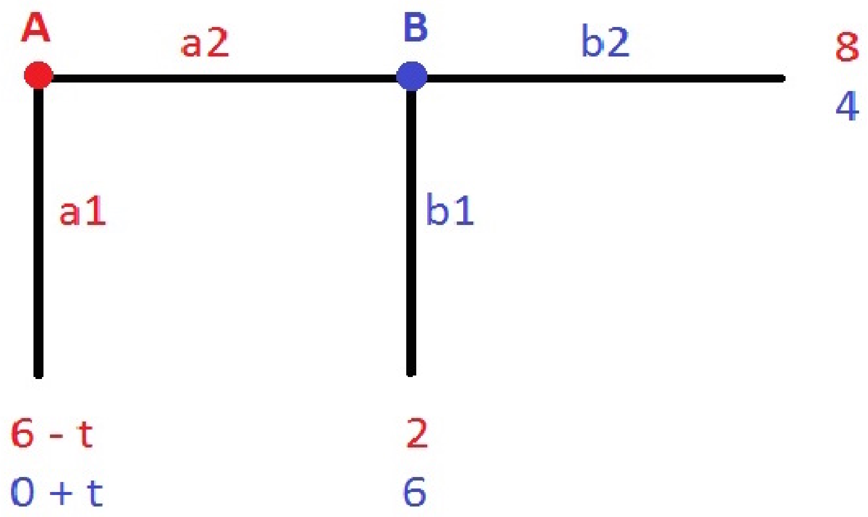

Figure 6: The Trust game

and A obtains 6 ECUs more—with respect to when he exploits trust. Thus, a player B motivated by social preferences of the kind that we elicited, i.e., distributional, may be willing to forego 2 ECUs in order to improve A's payoff of 6 ECUs. For what regards player A's payoffs, we implement the compensation t for the following reason. As in the CG, A may not pass money to B for insufficient trust and not for the lack of a social concern. Thus, subjects choosing a_1 do not necessarily reveal that they are selfish. Clearly, once strategic reasoning induces player A to play a_1, a positive transfer to B would reveal a social concern for B's payoff, of the distributional kind.

In the Trust game, 37.1% of our subjects chose b_2 in B's role, while 29.3% chose a_2 in A's role and, among those who chose a_1, 46.4% chose a transfer larger than zero.

In Figure 7 and 8 we report the estimated kernel density of the normalized SVO types (θ) by actions in the Trust game. An inspection of these graphs reveals that at the value 0.64 of the parameter θ there is a reversal of the relative likelihood of choices in the Trust game: actions b_2 (b_1), and actions a_2 (a_1) are more frequent above (below) this threshold. Moreover, probabilities of choosing b_2 (a_2) positively correlate with the normalized SVO angle (OLS estimation, 1% significance level).

Since A players may choose a_1 because of insufficient trust and not because they lack a social concern for the co-player's payoff, it is worth relating the choice of a positive compensation t to the SVO angle classification. Figure 9 reports the Kernel density of types choosing $t > 0$ (blue line) and of those choosing $t = 0$ (red line). The two

Yet, the literature on the Trust game has shown a certain flexibility in the choice of the payoff structure satisfying some restrictions that are in line with the features of the CG (see, e.g., Kreps et al., 1996, Anderhub et al., 2002, Wang and Ng, 2015): when trust is honored both players get more than what they get in the absence of trust, the trustee has a material incentive to exploit trust and in case he does the trustor gets less than what he would get from not trusting.

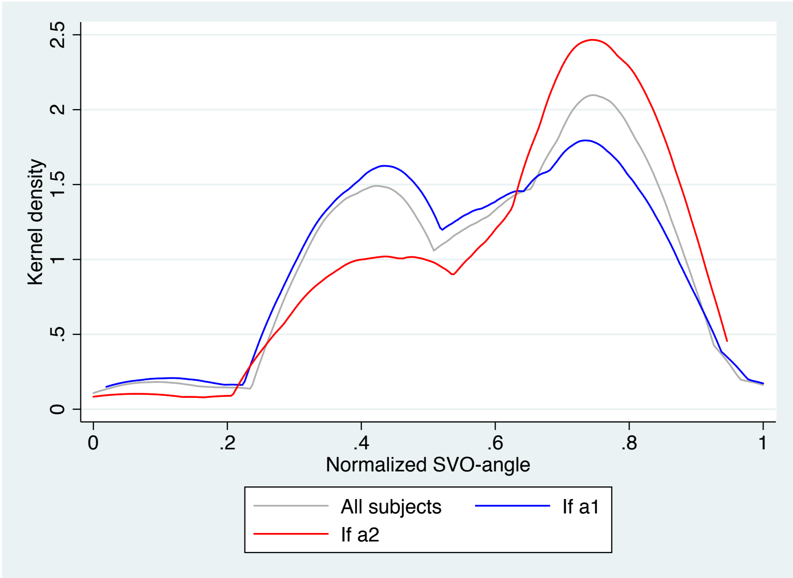

Figure 7: Kernel density estimation of the normalized SVO angle (θ) conditional on choosing a_1 (blue line) and a_2 (red line) in the Trust game; the grey line represents the unconditional Kernel density of types.

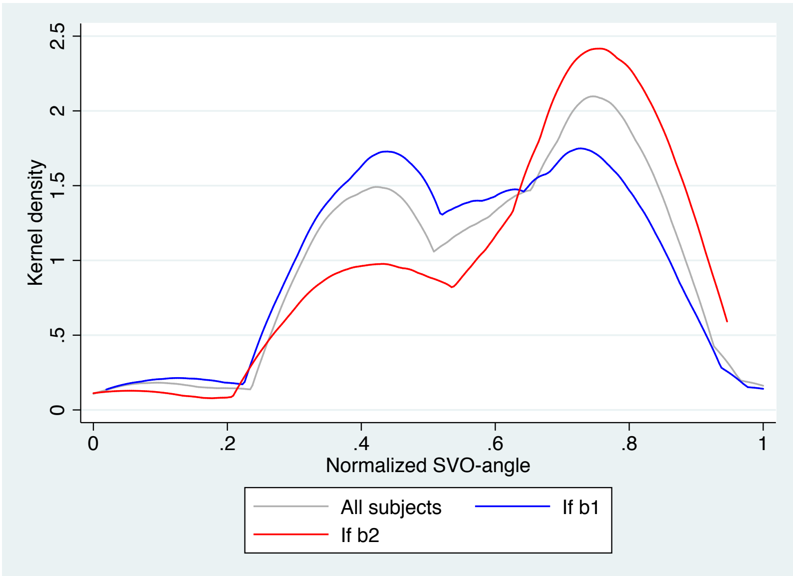

Figure 8: Kernel density estimation of the normalized SVO angle (θ) conditional on choosing b_1 (blue line) and b_2 (red line) in the Trust game; the grey line represents the unconditional Kernel density of types.

densities cross at a normalized SVO angle of 0.61. Moreover, the compensation and the normalized SVO angle are positively correlated (OLS estimation, 1% significance level).

Given these results, we conclude that a threshold of 0.59 of the SVO angle like the one proposed by Murphy et al. (2011) to discriminate between pro-selfs and pro-socials can meaningfully contribute to explain choices in the CG. As discussed in the main text (see Section 3.2), there is a mis-classification of 14 subjects between the threshold of Murphy et al. (2011) ($\bar{\theta}$=0.59) and the highest threshold derived from inspection of kernel densities ($\bar{\theta}$=0.64). However, the sensitivity analyses discussed in Section 4.4 show robustness of results to the choice of the threshold.

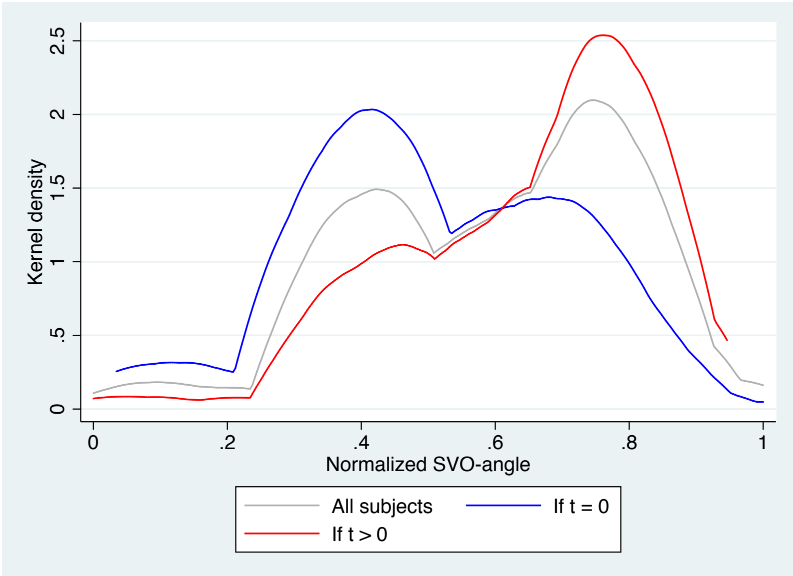

Figure 9: Kernel density estimation of the normalized SVO angle (θ) conditional on choosing a transfer $t = 0$ (blue line) and $t > 0$ (red line) in the Trust game (given a_1); the grey line represents the unconditional Kernel density of types.

C. Beliefs: elicitation instructions, distributions, and analysis

In this section we provide details about the beliefs elicitation procedure. Figures 10 displays the message announcing the beliefs elicitation in round 1. Analogous screens were displayed at rounds 17, 18, 19 and 40, but after play. Figure 11 and 12 display the screen where subjects had to insert their beliefs in the two methods of play. To ease participants' understanding of the task, these instructions were using examples of hypothetical beliefs (50 and 25) and frequencies (between 45% and 55% and between 20% and 30%). We checked that these specific examples did not induce any particular clustering of responses along the distribution of elicited beliefs. Below in Figure 13 and 14 we show these distribution in round 1 for the two methods. Analogous evidence characterizes the distributions at later rounds and is available on request.

Figure 10: Announcement of the beliefs elicitation in round 1

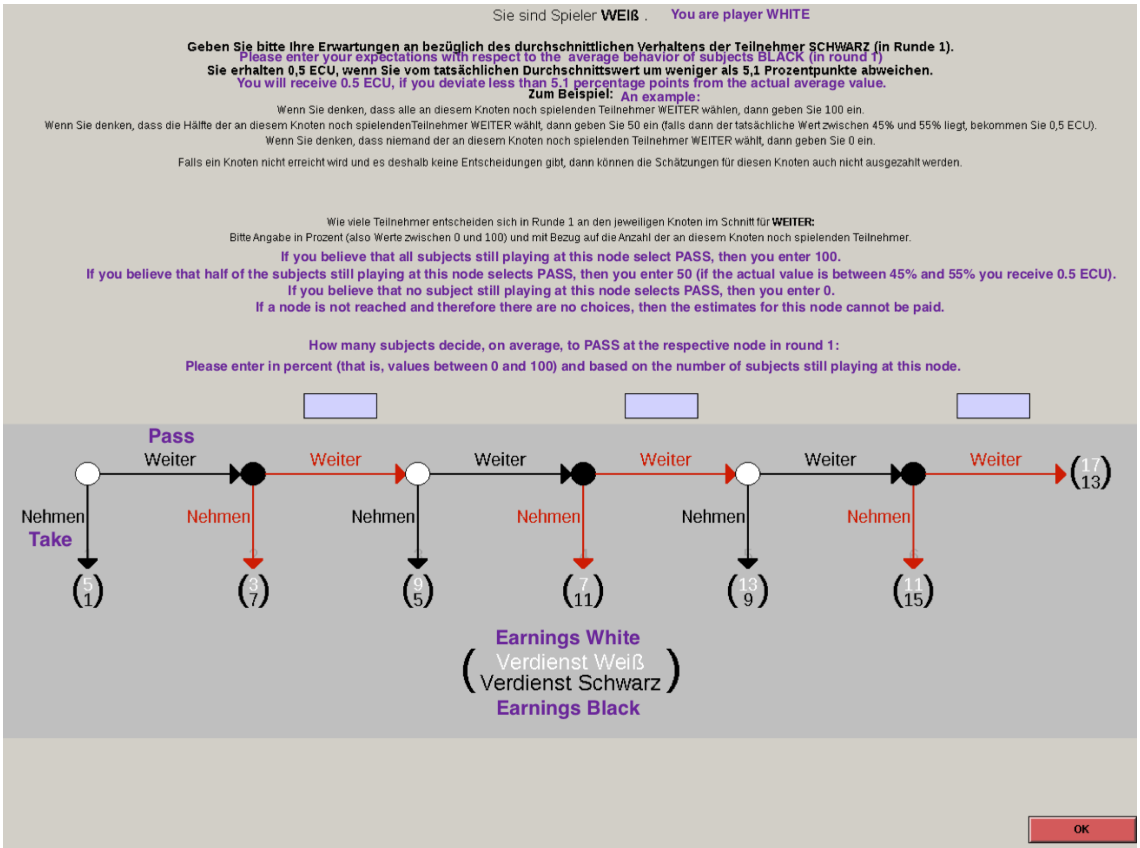

Figure 11: Beliefs elicitation page for White in Direct in round 1

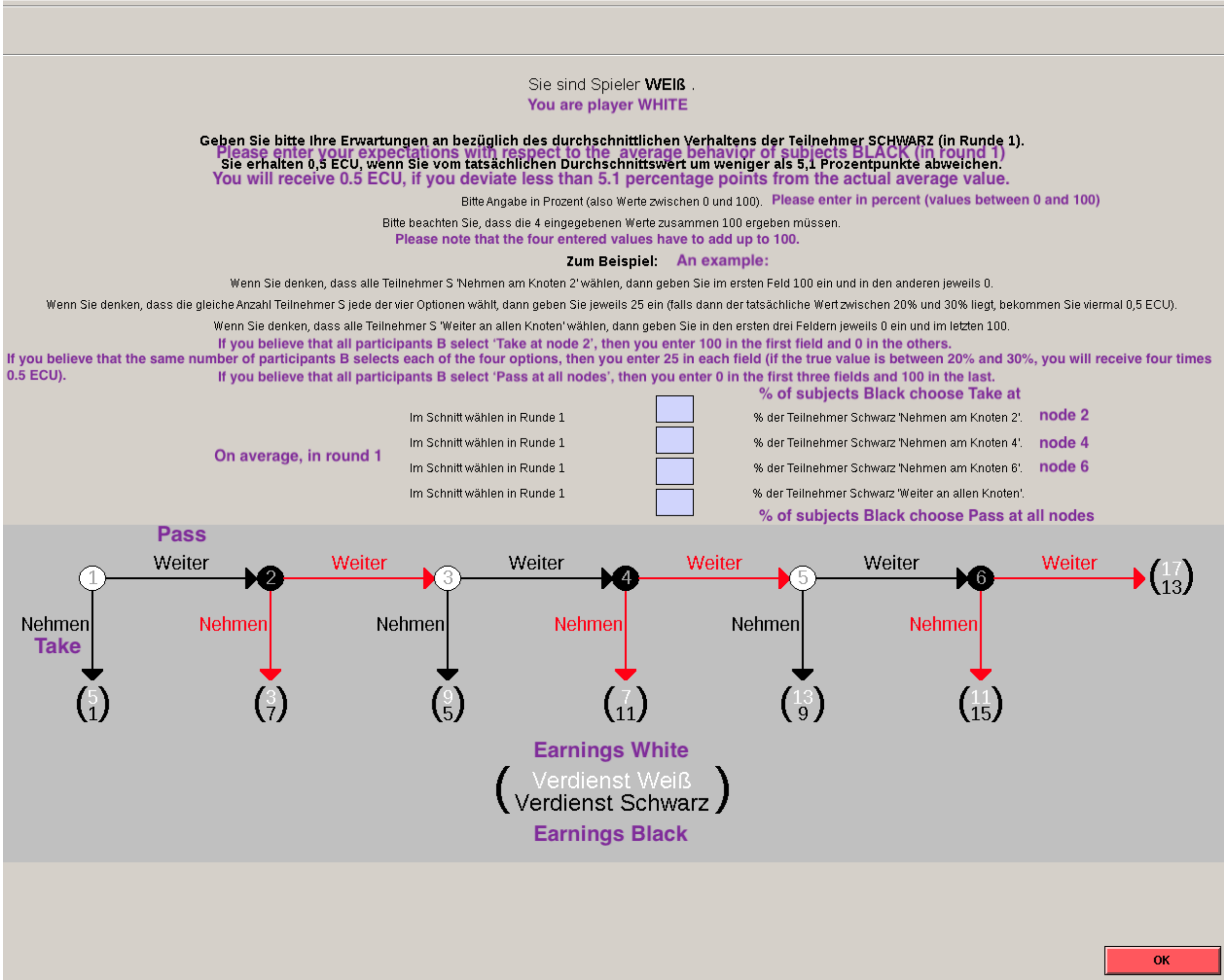

Figure 12: Beliefs elicitation page for White in Strategy

We provide further information on the outcome of beliefs elicitation by showing the aggregate distributions of beliefs, pooled across methods (after converting strategy data in the same metric of direct data). See Figures 15, 16 and 17 below. They show that within rounds beliefs at w_1 and b_1 are right-skewed, beliefs at w_2 and b_2 are roughly uniformly distributed, while beliefs at b_3 are left-skewed and this pattern tends to reinforce as rounds unfold.

Finally, in Table 8 we report regression results for the impacts of type and treatments (and their interactions) on aggregated middle rounds beliefs, with the same specification that we use to analyze round 40 data in the main text. Results show that in Public beliefs are consistently higher, except at w_1. The correlation between θ and beliefs at w_3 is positive and significant in Strategy. However, it is not considering the negative main effect of θ (Wald test, $p = 0.09$). Across all nodes of White pro-socials' beliefs in Personal-Strategy are higher than pro-selfs' ($p = 0.05$), while they are not significantly different in Public-Strategy ($p = 0.94$).

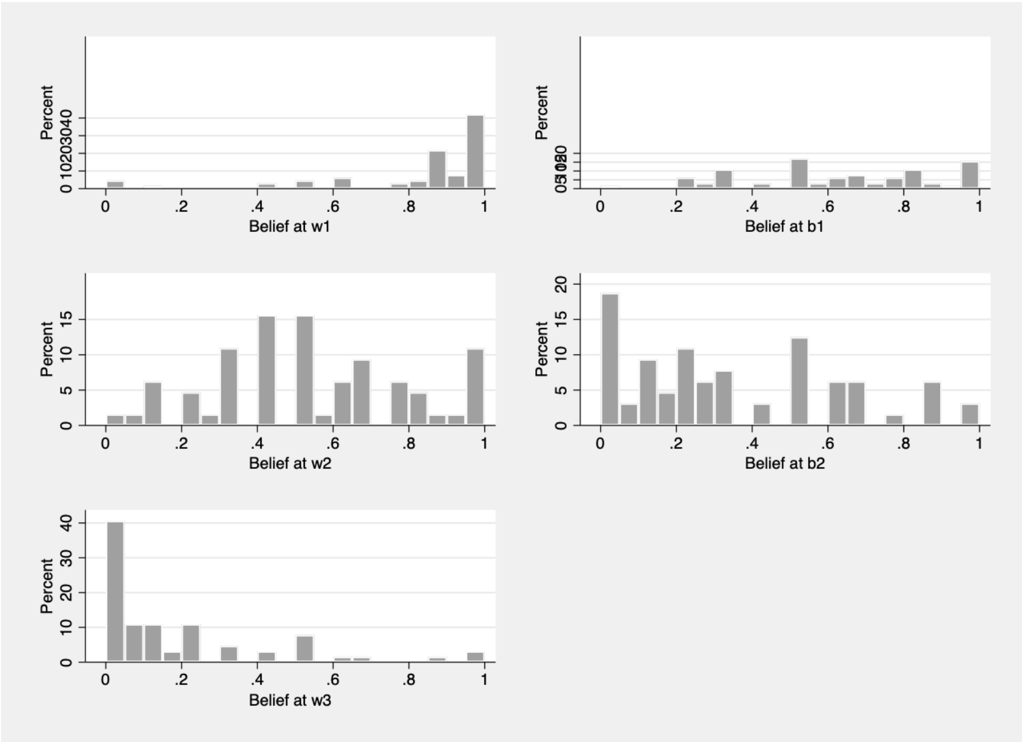

Figure 13: Histograms of the distribution of beliefs in Direct (round 1). Beliefs at one node are about whether the opponent will play Pass at the next node (e.g., the w_1 belief is about the choice at b_1, the b_1 belief is about the choice at w_2, etc.) This means beliefs data are not available for b_3.

Table 8: Round 17 to 19 beliefs

	w_1		b_1	
θ	-0.00016	(0.0016)	-0.0024*	(0.0013)
Public	0.054	(0.034)	0.18***	(0.043)
Strategy	0.023	(0.038)	0.0024	(0.041)
$\theta \times$ Public	0.00056	(0.0011)	0.0027*	(0.0016)
$\theta \times$ Strategy	-0.000019	(0.0014)	0.0039**	(0.0016)
$\theta \times$ Strategy $\times$ Public	-0.00064	(0.00098)	-0.0049***	(0.0013)
Constant	0.93***	(0.051)	0.80***	(0.033)
	w_2		b_2	
θ	0.0033	(0.0029)	-0.0010	(0.0021)
Public	0.40***	(0.061)	0.41***	(0.074)
Strategy	0.049	(0.064)	0.0062	(0.073)
$\theta \times$ Public	-0.0025	(0.0028)	0.0036	(0.0027)
$\theta \times$ Strategy	0.00084	(0.0031)	0.000047	(0.0030)
$\theta \times$ Strategy $\times$ Public	-0.0023	(0.0024)	-0.0032	(0.0020)
Constant	0.42***	(0.063)	0.25***	(0.052)
	w_3			
θ	-0.0017	(0.0016)		
Public	0.25***	(0.053)		
Strategy	-0.0064	(0.055)		
$\theta \times$ Public	0.0041**	(0.0018)		
$\theta \times$ Strategy	0.0049**	(0.0020)		
$\theta \times$ Strategy $\times$ Public	-0.0065***	(0.0017)		
Constant	0.14**	(0.055)		
ρ_{12}	.295 ***	(0.07)	0.265 ***	(0.062)
ρ_{13}	-0.101	(0.077)	–	–
ρ_{23}	0.259 **	(0.099)	–	–
Observations	384		384	

Notes: System of seemingly unrelated regressions with sequential partial observability; the dependent variables are the beliefs in rounds 17-19 at each decision node of White (Black) about the choice of Black (White) at the next node; robust standard errors in parentheses; significance levels are: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

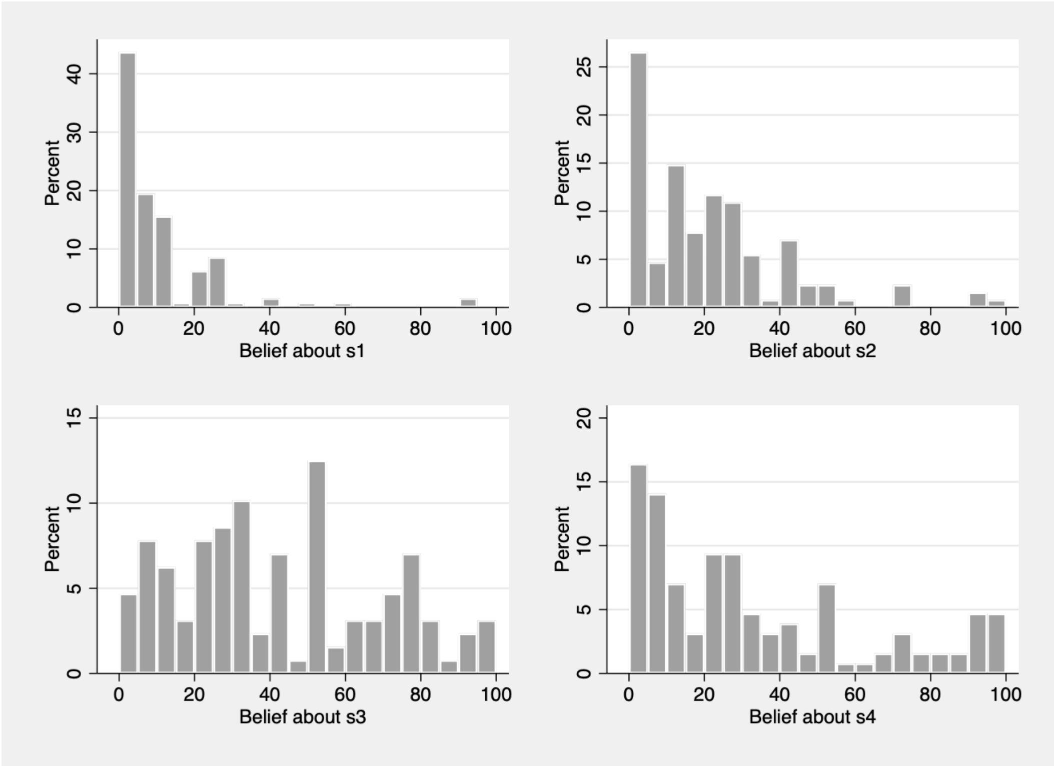

Figure 14: Histograms of the distribution of beliefs in Strategy (round 1). Beliefs are about where the opponent will choose strategy s_1, s_2, s_3 or s_4.

D. Background on the payment mechanism

Azrieli et al. (2018) show that in experiments with a sequence of games, pay-all is the incentive compatible mechanism when there are no complementarities that may reverse preferences over plans of actions in a game. Yet, when preferences display complementarities (e.g. wealth effects, spillovers, overall fairness considerations) pay-all is not incentive compatible. We cannot fully exclude that our setting (Trust game and the CG) is free of complementarities. However, due to the pronounced asymmetry in the two tasks—one round in the Trust game versus 40 rounds in the CG—we considered the impact of cross-task complementarities to be minor in comparison to a potential dilution of incentives in the CG, which is the focus of our study, if only one random task were paid. We doubt that any across-task consideration subjects could make would persist as the CG rounds unfold.

Subjects knew that the payments from each part will be announced at the very end of the experiment. While this approach might exacerbate hedging, we decided that avoiding contamination of CG choices or beliefs is more important: being informed about payoffs of the Trust game, a strategic interaction similar to the CG, might lead to a beliefs update about opponents' possible behavior in the CG.

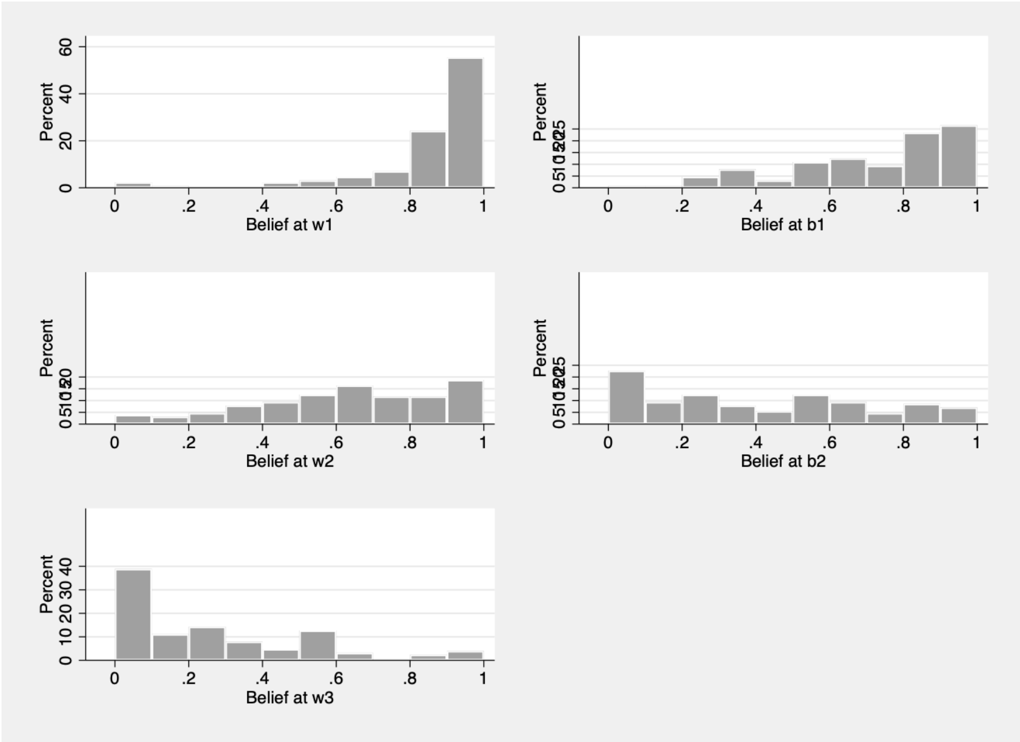

Figure 15: Distribution of beliefs in round 1 (all treatments). Beliefs at one node are about whether the opponent will play Pass at the next node (e.g., the w_1 belief is about the choice at b_1, the b_1 belief is about the choice at w_2, etc.). This means beliefs data are not available for b_3.

Hedging is also a possibility between incentivized beliefs and choices (see Schotter and Trevino, 2014, for a review of proper scoring rules for belief elicitation). Commonly, paying relatively small amounts for accurate beliefs is implemented to reduce incentives to hedge (Rutström and Wilcox, 2009, Blanco et al., 2010, Armantier and Treich, 2013). While this approach does not guarantee eliminating hedging incentives, we decided to follow the conventional approach by keeping beliefs payoffs quite small as a means to minimize the scope for hedging.

When considering incentives within the CG, other concerns arise. Azrieli et al. (2018) argue that the optimal incentive mechanism for multiple rounds strategic interaction experiments depends on whether the environment is a repeated game (same opponent) and on whether subjects can influence through their behavior the feedback they get on opponents' play. While our 40 rounds CG is not a repeated game, feedback in Personal is endogenous to own play. This feature creates scope for experimentation, inhibiting truthful play: subjects may choose to Pass at some information set in a round, even it is the less preferred option, in order to acquire valuable information about opponents' behavior and make a more profitable choice later on. Thus, a pay-all mechanism would best fit the Personal information treatments.

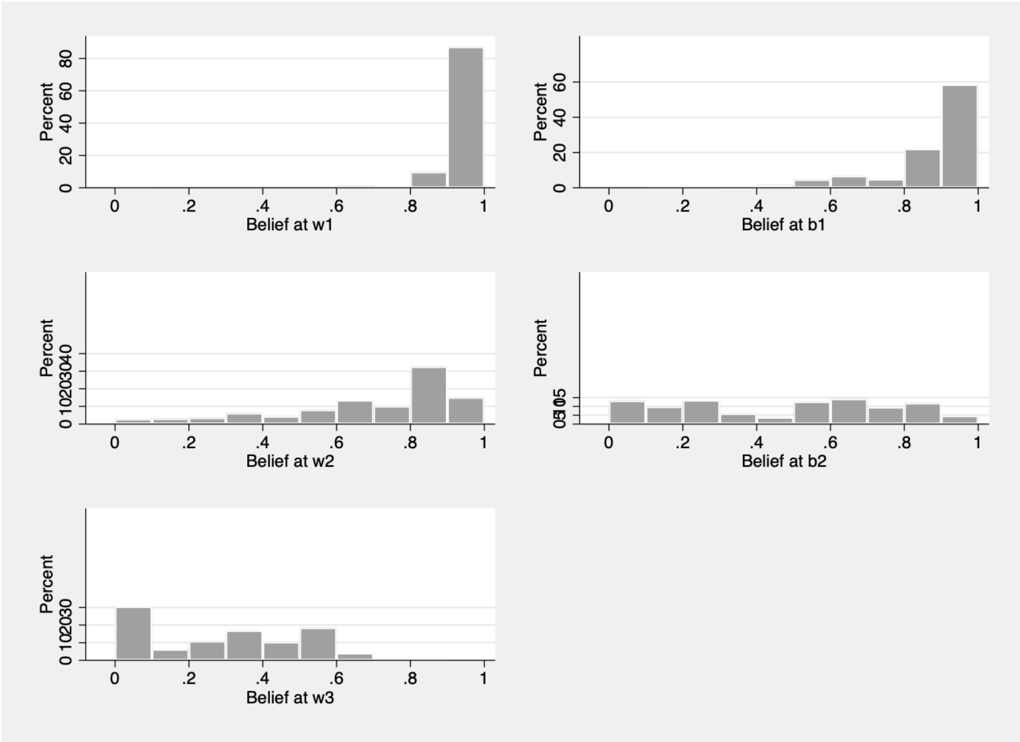

Figure 16: Distribution of beliefs in round 17 to 19 (all treatments). Beliefs at one node are about whether the opponent will Pass at the next node. For example, w_1 is the belief about whether the opponent will Pass at b_1. Note that beliefs data are not available for b_3.

Instead, a pay-one mechanism would better fit the Public information treatments, where feedback is exogenous to own play and there is less scope for experimentation. Since we cannot implement different payment schemes across treatments, we opted against the pay-all mechanism. This choice has also the advantage of avoiding wealth effects. Yet, in order to limit the scope for experimentation that may affect play in Personal, we decided to pay one round in the first half, where incentives to experiment are stronger, and one in the second half.[44]

[44] In the closely related literature, Danz et al. (2016) opt for the same payment mechanism as we do, while Nagel and Tang (1998) chose a pay-all mechanism.

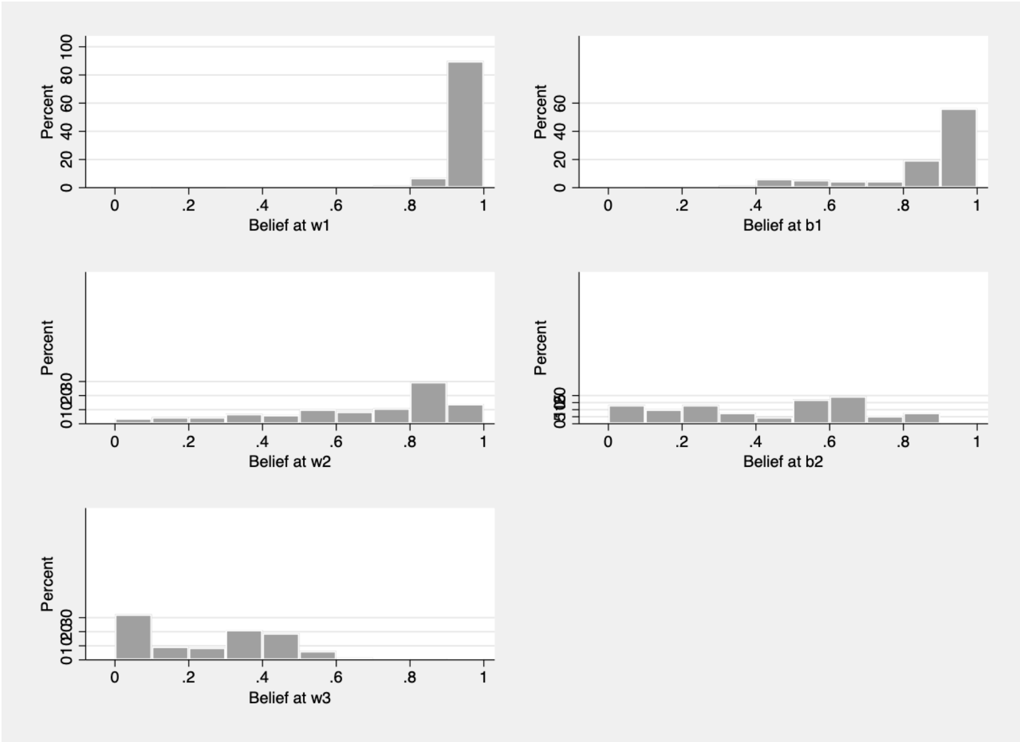

Figure 17: Distribution of beliefs in round 40 (all treatments). Beliefs at one node are about whether the opponent will Pass at the next node. For example, w_1 is the belief about whether the opponent will Pass at b_1. Note that beliefs data are not available for b_3.

E. Treatments comparison for Strategy only

Table 9: Information treatments comparison (only Strategy)

	Whites		Blacks	
θ	0.059**	(0.023)	0.0687***	(0.024)
Public	2.91***	(0.717)	2.779***	(0.674)
$\theta \times$ Public	-0.069**	(0.030)	-0.070**	(0.030)
Observations	640		640	

Notes: Ordered probit regressions by role for rounds 31-40; the dependent variable is Choice which takes value 1, 2, 3 and 4 when the strategy chosen is s_1, s_2, s_3 and s_4, respectively; robust standard errors in parentheses; significance levels are: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.