Experimental Economics (2022) 25:1173–1201
https://doi.org/10.1007/s10683-022-09750-1

ORIGINAL PAPER

The Influence of Indirect Democracy and Leadership Choice on Cooperation

Fanny E. Schories[1]

Received: 6 November 2020 / Revised: 6 February 2022 / Accepted: 22 February 2022 /
Published online: 15 April 2022

Abstract

The paper examines whether an institution has a differing impact on cooperation if it is introduced by a representative of the affected subjects rather than exogenously imposed. The experimental design controls for selection effects arising from the endogenous policy choice. The treatment varies whether the decision-maker is elected or randomly appointed. There is evidence of a large democracy premium in the sense that endogenously chosen institutions lead to more cooperation than identical exogenous institutions, but only if the group leader is democratically chosen. Especially the subjects who initially did not prefer the policy are more likely to cooperate if it was brought about by an elected representative. There is no democracy premium for randomly appointed group leaders.

Keywords Laboratory experiment · Representative democracy · Collective decision-making · Social dilemma · Legitimacy

JEL Classification C9 · D02 · D72

A part of the paper was previously circulated as a working paper under the title "Institutional Choice and Cooperation in Representative Democracies: An Experimental Approach". I would like to thank the editor, Lata Gangadharan, two anonymous referees as well as James Andreoni, Pedro Dal Bó, Felix Hadwiger, Julia F. Körner, Shushanik Margaryan, Lydia Mechtenberg, Jana Mintenig, Gerd Mühlheußer, Louis Putterman, and especially Julia Wolffson. This work was financially supported by the Graduate School of Economics and Social Sciences and the DFG Graduate College in International Law and Economics, both at Universität Hamburg. Replication materials, including data and instructions are available at https://github.com/Schoriesfanny/leadership-choice-cooperation.

✉ Fanny E. Schories
fanny.schories@ile-graduateschool.de

[1] Department of Economics, Hamburg University, Hamburg, Germany

1 Introduction

Does the way a law is implemented influence its effectiveness? An experiment is used to quantify the behavioral difference between externally imposed institutions and those that are implemented through a democratic procedure. This behavioral difference has been coined *democracy premium*: When holding information and group composition constant, there is an increased willingness to cooperate after an institution was introduced through a democratic procedure.

That a central authority can effectively improve cooperation is a stylized fact in experimental economics. However, most studies take these institutions as exogenously given. Recently, many experiments studied endogenous institutions which do not "fall from heaven" but are introduced by the affected parties themselves. The process that leads to an institutional setting might well influence to what extent it can fulfill its societal purpose. Should a law that was effective in one instance also be assigned in other situations? Consider the progressing European integration as an example. Is a reform introduced in a member state as effective when it is de facto prescribed from an external authority such as the "Troika" as if it was introduced autonomously by the elected national government?

Previous studies found inconclusive evidence regarding the existence of a democracy premium in direct democracies (Dal Bó et al., 2010; Sutter et al., 2010; Vollan et al., 2017; Gallier, 2020). This study focuses on representation as another aspect of democratic decision-making. Indirect democratic procedures are commonly used: nations, firms, and clubs typically delegate at least parts of their decision-making processes to representatives. The paper contributes to two central questions: Is there a democracy premium when a representative chooses the relevant institution? And does the democracy premium depend on how the representative came into office? The two experimental treatments address the second question. The treatment variable is the way the group leader is appointed: via election or lottery. The treatments are therefore called indirect democracy (ID) and random dictator (RD). Indirect–or representative–democracy is the most common form of democracy today (Alizada et al., 2021). Using lots to assign political roles, also known as *sortition*, dates back to Athenian democracy, where it was promoted by Aristotle to achieve the democratic ideals of equality and fairness more effectively than elections (Barnes (1984)).

The experiment has three stages. The first stage of the experiment consists of a prisoners' dilemma played in small groups. In the second stage, subjects form preferences about a payoff modification for their group in the final stage. The modification transforms the prisoners' dilemma into a coordination game, which makes both defection as well as cooperation incentive-compatible. In the indirect democracy treatment, subjects elect a group representative, whereas the group leader is determined by chance in the random dictator treatment. The leader's preference about changing the payoffs becomes binding for the group but is only considered in 50 percent of the cases. If it is not considered, one of the two games is randomly assigned to each group for the third stage. This design feature–first introduced by Dal Bó et al. (2010) (hereafter: DFP) – can control for information

and selection effects and thus allows a clean estimate of the democracy premium. Cooperation rates before and after the vote are analyzed conditional on individual policy preferences and the outcome of the random intervention to estimate the democracy premium for both treatments separately.

I find that the payoff modification significantly increases cooperation and even more so when introduced by an elected representative. Moreover, the difference between the treatments is striking. There is a substantial effect of endogenous choice in the representative democracy: 78 percent of the increase in cooperation can neither be attributed to the payoff change itself nor to differences in group composition and thus remains as the democracy premium. In contrast to DFP (2010) especially those who initially do not prefer the coordination game respond strongly to a democratic payoff modification with an increased willingness to cooperate. Conversely, the randomly appointed leader in the second treatment does not cause an increase in cooperation beyond the exogenous payoff modification. The effect of the random dictator is negative because cooperation is further decreased in the case of endogenous non-modification. The results have important implications for policy-making but also for the methods used in experimental economics.

The remainder of the paper is structured as follows. Section 2 reviews the related literature focusing on the effects of endogenous formal institutions in economic laboratory experiments. Section 3 presents the design of the experiment, including testable hypotheses. The analysis and results are presented in Sect. 4. Finally, Sect. 5 discusses potential explanatory approaches from economic theory and concludes.

2 Related literature

The focus of this paper lies on formal institutions that are exogenously enforced in the form of law or other regulation, as opposed to informal sanctions, which are maintained privately. Elinor Ostrom laid the groundwork for the experimental study of self-governance as a way of overcoming collective action problems (e.g. Ostrom (1991, 1992)). By now, there is a large and growing body of experimental literature exploring the key factors that influence cooperative behavior in societies. These studies suggest that the implementation of an institution matters in addition to the institutional design itself. A central result from previous lab experiments on such institutions is that direct democratic participation rights increase subjects' contributions to a public good, ceteris paribus.[1]

[1] Various authors have taken the search for effects of endogenous institutions in public goods games to the field. Cavalcanti et al. (2010) find that public deliberation increases the willingness to contribute to projects for the management of common resources among Brazilian fishermen. Other studies such as Bonin et al. (1993), Bardhan (2000), Black and Lynch (2001), and Fearon et al. (2011) find similar results in settings ranging from irrigation rules in rural India to workplace decisions of manufacturing businesses in the USA: participation rights increase compliance, productivity, and satisfaction. Grossman and Baldassarri (2012) find that subjects electing leaders contribute more to public goods than subjects who were assigned leaders through a lottery.

The evidence about the influence of participation rights on cooperation levels is based mainly on public goods experiments, where democratic structures are implemented into the policy selection process by allowing participants to vote on different proposals directly (see Chaudhuri (2011) for a survey of the literature on cooperative behavior in public goods games and Bó and Pedro (2014) and Dannenberg and Gallier (2020) for the literature on endogenous institutions). Tyran and Feld (2006) show that an endogenously chosen non-deterrent law reduces free-riding behavior. The experiment varies the severity and enaction of a monetary punishment on free-riding. An exogenously implemented mild law does not significantly increase compliance compared to the game without law. In the endogenous treatment, individuals mostly vote in favor of the mild law, and the contribution rate is significantly higher than without law (Tyran and Feld, 2006). Sutter et al. (2010) present additional evidence that participation rights reinforce cooperation. Subjects vote for a decentral punishment or reward mechanism. The endogenous choice is associated with higher contributions for any given institution compared to an identical mechanism implemented through an external authority (Sutter et al., 2010). Markussen et al. (2014), Kamei et al. (2015), Dannenberg et al. (2020) find that endogenously chosen sanctioning mechanisms improve public good provision. However, the experiments presented so far in this section cannot isolate a pure democracy premium. The vote entails a signaling component and reveals information about the group composition and subjects' preferences in the endogenous case. Conditionally cooperative players are likely to respond to this signal and adjust their behavior accordingly. Furthermore, because of the democratic policy selection, the institution is not randomly assigned, and the estimated differences between exogenous and endogenous assignment are potentially biased by self-selection. Both self-selection and information as confounding factors are mitigated using the experimental design presented in Sect. 3.1.

The experimental mechanism employed in the present study was first introduced by DFP (2010). It avoids a self-selection bias in experiments investigating the effect of democracy. Subjects are allowed to vote on a fixed policy proposal, but this democratic choice is overruled by a random computer decision in 50 percent of the cases. This strategy makes it possible to control for unobservable characteristics that influence both voting decisions and cooperative behavior. DFP (2010) find that even when controlling for selection, endogenous and exogenous institutions have a differing impact on cooperation. The authors find evidence of a democracy premium: a cooperation-enhancing influence of democratic institutions beyond the instrumental effect of the policy choice. Sutter et al. (2010) obtain contradictory results using a similar randomization mechanism: whether the vote was considered has no significant influence. The authors conclude that the institutional design itself influences behavior and not its democratic implementation. Kamei (2016) follows the randomization mechanism suggested by DFP (2010) and finds not only a direct democracy premium but also pro-social behavioral spillovers from democratic procedures: Those involved in an endogenous institution formation keep their increased cooperativeness even in a following setting without democracy. Gallier (2020) combines the experimental techniques used in Tyran and Feld (2006) and DFP (2010). He finds ambiguous results regarding the democracy premium: overall, contributions to a

public good are higher if a sanctioning institution is democratically implemented. However, the difference is driven mainly by self-selection and information effects, which the identification strategy can precisely estimate. Only for the subjects that initially did not prefer the institution does an actual democracy premium exist. Vollan et al. (2017) find the opposite of a democracy premium in China. Instead, the Chinese subjects are more prone to comply with an exogenous policy. The authors attribute this result to a culture of obedience towards authority. No positive democracy premium has been found in a representative democracy yet (Castillo et al. 2017; Kamei, 2017).

The treatment variation presented in this paper directly links to the strand of literature on the differences between elected and appointed leaders. A regularity already established in the literature is that elected leaders behave more pro-socially than randomly appointed ones for several reasons, e.g., because the election functions as a screening tool, because it creates accountability or because the leaders demonstrate reciprocity towards voters (Brandts et al., 2006; Hamman et al., 2011; Corazzini et al., 2014; Brandts et al., 2015; Marcin et al., 2019; Drazen and Ozbay, 2019). The threat of reelection is not present in the paper here, such that the design is most closely related to Drazen and Ozbay (2019), who compare the behavior of elected and non-elected leaders in a citizen-candidate model setting and find that only the elected representatives demonstrate reciprocity towards the citizen. What differentiates this paper from the previously cited studies is that here the leader takes only one decision on behalf of the other subjects but then continues to act as a regular group member. Furthermore, voters have been found to be more satisfied and sustain higher cooperation if they were involved in selecting a leader (Rivas and Sutter, 2011; Andreoni and Gee, 2012; Mechtenberg and Tyran, 2019). Section 3.3 returns to the findings from the literature presented in this section and how they inform the hypotheses for the experiment conducted in this study.

3 The experiment

3.1 Experimental design

The aim of the experiment is first to investigate whether a policy that a group leader actively chose induces more cooperation than the same policy implemented via an exogenous mechanism, and second whether in addition the way the leader making the endogenous choice comes into office influences subjects' willingness to cooperate. Both questions are answered using a between-subjects design. The first is investigated using the identification strategy developed by DFP (2010). For the second part of the research question, the treatment varies how the leader is chosen: through an indirect democratic process [ID treatment] or randomly appointed [RD treatment]. The experiment consists of three stages (see Fig. 1).

The games are based on a standard prisoners' dilemma, in which players can take one of two actions: cooperate and defect (see Table 1).[2] The prisoners' dilemma has

[2] To preserve neutral framing the actions are labeled A and B in the experiment's instructions.

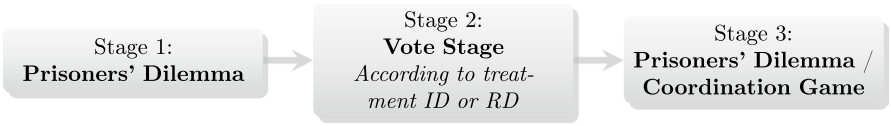

Fig. 1 Sequence of the Experiment

Table 1 Prisoners' Dilemma (left) and Coordination Game (right)

		Player 1			Player 2	
		A	B		A	B
Player 1	A	50, 50	30, 60	A	50, 50	30, 48
	B	60, 30	40, 40	B	48, 30	40, 40

a unique symmetric Nash equilibrium where both players defect. Following DFP (2010), ten rounds of the prisoners' dilemma are played in the first stage with random rematching of pairs within groups for every round. The groups are made up of four players and remain together over the entire session. When deciding on an action in the prisoners' dilemma, players do not know with whom they are paired. But after each round, as announced in the instructions, they are informed who their opponent was and what action each group member chose.

In the following vote stage, subjects make one or two decisions, depending on the treatment.[3] In both treatments, subjects first decide whether they want to change the payoff of their group to a coordination game for stage 3 or remain with their group in the prisoners' dilemma. The coordination game has a Pareto-superior Nash equilibrium in mutual cooperation (A,A). However, (B,B)–mutual defection–remains a Nash equilibrium in the coordination game as well. Every subject privately states a preference whether to implement the payoff modification or not. This decision will matter if she becomes the group leader. Second, the leader is determined according to the treatment. In the democracy treatment (ID), to elect the representative players privately announce another group member's player identification number without knowing her preference for modification. There is, however, complete information about the actions of every group member in the first ten rounds of the experiment. The player who is named most often in a group is elected as representative with plurality rule. Any tie is broken by the computer. In the random dictator treatment (RD), players do not vote for the representative. Instead, one player from each group is selected by a lottery placing the same probability on every subject. This way every small group has a leader, either randomly appointed or democratically elected, and this leader has stated a game choice for stage 3, which becomes binding for the group.

Analogous to DFP (2010), each leader's preferred game is implemented with a 50 percent probability. If it is not implemented, the computer chooses either the

[3] As noted by a referee, observed differences in behavior between the treatments could in principle also be triggered by the difference in the number of choices subjects make: one in RD, two in ID.

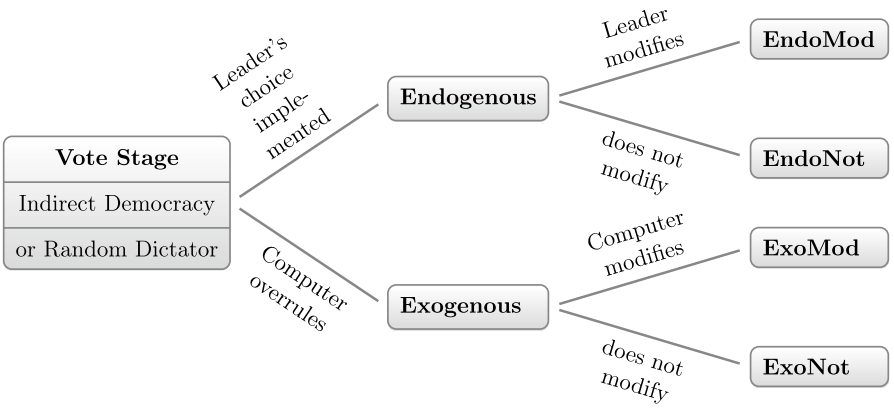

Fig. 2 Four Possible Vote Stage Outcomes (adapted from DFP (2010))

prisoners' dilemma or coordination game for the group, again with a 50 percent probability for each game. Consequently, there are four conditions under which subjects play stage 3 (see Fig. 2): payoffs modified to a coordination game by the group leader (EndoMod) or by the computer (ExoMod), and the unmodified prisoners' dilemma game either chosen exogenously (ExoNot) or by the leader (EndoNot). The twofold random intervention makes it possible to compare groups whose leaders decided in the same way but ended up in different conditions. The intervention is crucial to control for self-selection. Assuming there are unobservable player characteristics that increase both the preference for the coordination game as well as the willingness to cooperate, the treatment assignment is non-random whenever subjects choose their own payoff structure (by voting or otherwise). However, as DFP (2010) show, once the analysis conditions on an individual's modification preference and the implemented payoff structure, correlated unobservable characteristics are controlled for. After the vote stage, subjects are informed about the leader's player ID and game choice, whether the choice was considered, and the game the group will ultimately play in the last stage. Consequently, the experimental design controls for payoff modification preferences via the strategy method and holds information constant across outcomes. The respective game is then played for another ten rounds in stage 3.

3.2 Self-selection and information effects

Two aspects of the design are worthy of discussion, as they are inherently relevant to all experiments on endogenous institutions: self-selection and information effects. The main confounding factor of experiments investigating democracy is self-selection: whenever a policy is endogenously introduced, participants select into the treatment by definition. Naturally, players vote for the policy that is aligned with their preferences, making it difficult to compare the impact of different institutional designs since the assignment is not random. Cooperative subjects are more likely to prefer the policy that fosters cooperation. Self-selection thus leads to an

overestimation of the effect of endogenous policy selection since the voting decision and behavior are positively correlated (DFP, 2010). The randomization and elicitation of modification preference are introduced into the experiment to eliminate the self-selection effect from the analysis. The individual vote for or against the payoff modification can serve as a control for unobserved underlying characteristics influencing the willingness to cooperate. This design holds the advantage that the results of subjects that voted in the same way and ended up with the same game, but through a different mechanism, can be compared. DFP, (2010)'s identification strategy relies on the assumption that groups with an identical distribution of votes for and against modification also have identical preferences about modification and thus cooperation. If their behavior differs, this is attributed to the way the modification was implemented.

The second confounding factor that the design controls for is information. Sutter et al. (2010)'s as well as DFP (2010)'s main experiment involve instructing subjects about their group's choice only in the endogenous case, thereby straining the ceteris paribus assumption: not only does the institution's implementation differ between the endogenous and exogenous condition, but also the information provided. One can argue that this information is an essential part of a democratic institution, and the asymmetry between the outcomes should not be erased. But in this case it is not possible to truly isolate the democracy premium from information effects. Gallier (2020) systematically varies the informational content in the exogenous condition and finds that the information effect is a significant driver of heterogeneities in responses to the policy implementation. In the present experiment, subjects are thus informed about the leader's choice in the endogenous *and* in the exogenous outcomes. The design ensures that there are no differences in the amount or quality of available information between the endogenous and exogenous vote stage outcomes that subjects could condition their behavior on in the following rounds, except with regard to the policy implementation according to the treatment. If the leader's choice of game was revealed only in the endogenous condition, subjects would be able to update their beliefs about their group members in a way that the subjects in the exogenous case could not. Holding the available information constant across outcomes means that observed differences can be attributed to the intrinsic difference between democratically and exogenously introduced policies. I consider this a conservative estimate of the democracy premium because both above-mentioned related factors inherent to democratic processes are controlled for.

A potential limitation of the experimental design is that the identification strategy requires making use of the strategy method instead of a direct response mechanism. Because all subjects decide on their preferred policy before learning whether they became group leader and whether the decision was considered, the behavioral effect of the endogenous condition may be somewhat diluted. In a "hot" decision environment, the response to the decision-making procedure could be expected to be more pronounced.[4] Moreover, all vote stage outcomes, including the endogenous ones,

[4] There is mixed evidence on the influence of the elicitation method on behavior in experiments. See e.g., Brandts and Charness (2000); Brosig et al. (2003); Fischbacher et al. (2012)

could be perceived as fundamentally exogenous, since in each case the computer first decided whether to override the leader's decision or not. However, these factors can be expected to lead to an underestimation of the "true" democracy premium, if they matter at all.

3.3 Hypotheses

Following previous literature and the discussion in this section so far, we can formulate specific hypotheses to be tested by the experiment. Unless explicitly stated, all hypotheses apply to both treatments and are tested separately and between subjects.

It is a weakly dominant strategy for subjects in both treatments to choose the game they truly prefer, a necessary assumption for the identification strategy. The individual game choice never influences who becomes the group leader: the game preference of a subject remains private information during the election in ID, and leaders are chosen randomly in RD. Hence, there is no incentive to misrepresent game preferences for the sake of becoming group leader, and the game choice is strategy-proof in that regard. Furthermore, the optimal choice between the games does not depend on the treatment.

Which game is preferred depends on a subject's beliefs about the other group members' actions in the coordination game and their own player type. Following previous literature on the public goods game – which presents a social dilemma situation like the prisoners' dilemma – as surveyed by Chaudhuri (2011), we can assume there to be three player types: unconditional cooperators who always play A, conditional cooperators who prefer to play A if they believe their opponent will do the same and B otherwise, and defectors who play B regardless. Most studies found that the majority of subjects are conditional cooperators. It is thus a reasonable assumption for subjects of all three types to believe that they are paired with conditional cooperators. In the prisoners' dilemma, the behavior of conditional cooperators will depend on their actual group composition: if there are other (un-)conditional cooperators in the group, the outcome (A,A) may be played, even though there is the constant temptation to deviate to the dominant strategy of playing B, which the defectors certainly choose. If a conditional cooperator is paired with one or more defectors, she will reciprocate by playing B as well.[5] Therefore, the likelihood of observing cooperative behavior in stage 1 decreases disproportionately in the number of defectors in each group. As DFP (2010) note, whether or not a subject prefers to modify the payoffs for stage 3 depends on what equilibrium they expect to coordinate on in the coordination game. A defector is weakly better off in the prisoners' dilemma: since she always plays B, in equilibrium she will have the same payoff in both games, but can expect a higher gain off equilibrium should she meet a cooperating subject. On the contrary, all conditionally and unconditionally cooperative

[5] As long as a conditional cooperator believes the probability of her opponent to play A is at least $\frac{5}{6}$, she should cooperate in the coordination game.

subjects at least weakly prefer the coordination game over the prisoners' dilemma as it turns mutual cooperation into an equilibrium.

Hypothesis 1 A subject who is cooperative in stage 1 is more likely to have a preference for payoff modification than an uncooperative subject.

The experiment is able to test three factors influencing cooperative behavior that are identified in the literature on the effects of elected leadership: selection of more cooperative leaders, intrinsic motivational changes of elected leaders', and changes in voter behavior in response to the decision-making process (Drazen and Ozbay, 2019). While the emphasis is on the latter, which includes the democracy premium phenomenon, the former two aspects are also testable using the experimental design at hand.

First, according to the selection factor, it is to be expected that groups in ID elect the most cooperative subject as representative (Hamman et al., 2011). It follows directly from Hypothesis 1 that the highly cooperative subjects should vote for another cooperative subject. Subjects who are conditional cooperators and play equilibrium strategies in which they defect in the prisoners' dilemma but cooperate in the coordination game are better off under the modified payoffs and should thus vote for a more cooperative subject as well. Only subjects who always defect can have a weak preference for another non-cooperative subject as representative under the assumption that she will not modify payoffs and that there are some other subjects in the group who cooperate off the equilibrium path. In short, only off-equilibrium beliefs justify having strong preferences for either the unmodified payoffs or an uncooperative representative.

Hypothesis 2 In treatment ID, cooperative players are more likely to be elected as representatives.

Second, elected leaders have been shown to act less selfishly than randomly chosen leaders (Drazen and Ozbay, 2019). This can be tested via a comparison between the two treatments.

Hypothesis 3 In stage 3, leaders in ID behave more cooperatively than leaders in RD.

Third, the response of the voters to the implementation process is at the heart of the paper. Rational choice would predict subjects to be indifferent between the decision-making procedures as long as the outcome, including the information revealed about other players' preferences, remains constant. However, as discussed in Sect. 2, numerous empirical studies have rejected this null hypothesis in the case of direct democratic decision-making. We could therefore expect endogenous policy selection to lead to more cooperation than the exogenously imposed policy in both treatments (between-subjects comparison of cooperation rates in stage 3). Regarding the difference between direct and representative democracy, the direct process

seems to be seen as carrying higher legitimacy (Olken, 2010; Towfigh et al., 2016). Therefore, we can expect the democracy premium to be smaller in the present setting compared to DFP, (2010).

Hypothesis 4 There is a democracy premium, i.e. cooperation rates are higher when a policy is democratically introduced.

Finally, the treatments ID and RD serve to improve the understanding of the sources of the democracy premium. Two candidate transmission channels shaping compliance with the outcome of a decision-making procedure are legitimacy and authority. On the one hand, legitimacy can be narrowly defined as being derived from the consent of the governed through elections (Locke, 1983). More specifically, institutions which are responsive to citizens by allowing participation in the decision-making process carry *input legitimacy* (Schmidt, 2013). On the other hand, authority is concentrated political power, regardless of its source. If the two coincide, a feeling of mutual responsibility arises. This is the case in a representative democracy, where authority stems from the figure of the representative herself and legitimacy from her democratic election. A random dictator has authority but no legitimacy from an election process. Compliance is described as a crucial consequence of political legitimacy (Ham et al., 2017). Thus, depending on which treatment corresponds to higher cooperation rates following the decision-making process, we can isolate the more important transmission channel of the democracy premium between the two factors authority and legitimacy, leading to two competing hypotheses.

Hypothesis 5.a (Legitimacy) The democracy premium is larger in ID than RD

Hypothesis 5.b (Authority) There is no difference between the democracy premia in ID and RD.

4 Analysis

4.1 Protocol and summary statistics

12 sessions took place at Hamburg University between 2016 and 2019 with a total of 280 participants (140 per treatment). No subject participated in more than one session or treatment. Upon arrival to the lab, subjects were randomly assigned to a computer cubicle and received the instructions in written form.[6] At the end of a session, participants filled out an unpaid socio-economic questionnaire including

[6] A translation of the German instructions can be found in the Electronic Supplementary Material. Instructions were read aloud and every subject correctly answered a set of control questions to ensure the instructions were well understood. Instructions for the second and third stage were handed out after the end of the first stage in order not to influence behavior prior to the vote.

Table 2 Summary Statistics

	Min	Max	Mean	Std. deviation
Age	17	57	25.7	5.1
Correct logic questions	0	3	1.7	1.2
Payout (€)	6	12	8.8	1.2
Share				
Female	59 %			
Full-time student	97 %			
Economics student	38 %			
Game theory knowledge	29%			

Sample size is n = 280 with 140 subjects in each treatment

Frederick (2005)'s cognitive reflection test, which has been shown to be strongly correlated with strategic sophistication (Carpenter et al., 2013).

Table 2 presents summary statistics. 106 subjects identified themselves as male, 166 as female, and 8 subjects chose the option "other or prefer not to say". The laboratory used the software hroot (Bock et al., 2014) to invite subjects from a pool of around 7000 registered participants recruited on the main campus at Hamburg University. Hence, the vast majority of subjects are full-time students. Out of those, around one-third of students were economics or business majors. More than a quarter of the subjects stated that they had at some point taken a class in game theory. More than one third of the subjects answered all three of the cognitive reflection test's logic questions correctly. The payment was made according to the outcome of two randomly chosen rounds, one from the first and one from the last stage, with an exchange rate of 10 points = €1. Subjects earned €9 on average which is in line with the mean hourly wage of €10 that the lab promises since all sessions lasted less than an hour. Subjects privately collected their payment in cash at the end of a session.

4.2 Individual analysis

The following sections examine the experimental data in the light of the hypotheses formulated in the previous section. Unless indicated otherwise, p-values are derived from non-parametric statistical tests, i.e. the Wilcoxon rank-sum test for within- and the Mann-Whitney-U-test for between-subjects comparisons.

Average cooperation in the first stage, in which all groups played the regular prisoners' dilemma, amounts to 32 percent in the ID treatment and 36 percent in RD (see left panel of Fig. 3). The difference is small and not statistically significant. The general pattern of positive but decreasing cooperation is well-known from previous experiments using the prisoners' dilemma (Cooper et al., 1996).

In the vote stage, 64 percent of subjects in ID and 65 percent in RD chose to modify the payoffs in case they became group leader ($p = 0.803$). Table 3 shows linear probability models (LPM) and Probit estimates of voting for modification regressed on variables covering experiences from stage 1 and personal characteristics. Own cooperation is positively related to a preference for the coordination game

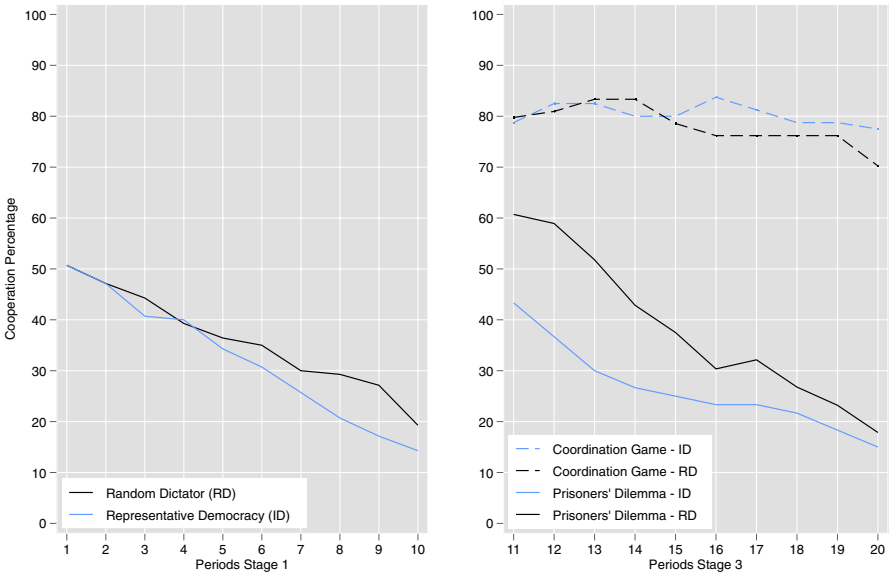

Fig. 3 Cooperation Rates in stages 1 and 3. Depicted are the shares of subjects choosing to cooperate in each round, separated by treatment (RD and ID) and game. In stage 1, all subjects play the prisoners' dilemma

and significant in every specification. A player who cooperated in all of the first ten rounds has a probability to favor the coordination game that is almost 50 percentage points higher than that of someone who did not cooperate at all. Furthermore, partners' cooperation has a significantly negative influence, which is intuitive since a subject with cooperative group members in stage 1 would see less necessity to switch to the coordination game to increase cooperation. Cognitive reflection, as measured by the variable *logic*, is significantly positively correlated with a preference for modification. A subject who was able to answer all three cognitive reflection questions correctly has a probability of voting for modification that is 30 percentage points higher compared to one who gave no correct answer. The treatment dummy variable *random dictator* is not significant, which supports the conjecture that modification preferences are formed independent of treatment. Overall, the data lends support to Hypothesis 1, implying that cooperative actors self-select into matching institutions.

Result 1 Cooperative individuals and those with a higher cognitive ability have an increased preference for the coordination game.

In the representative democracy treatment, nearly half of the participants (43 percent) declared in the questionnaire they had voted for a group member because it had appeared cooperative in the first stage. But in fact the elected representatives cooperated *less* than the other players in stage 1 (on average 2.5 versus 3.5 out of 10 rounds, $p = 0.022$). There is neither a significant difference in

Table 3 Individual Determinants of Institutional Preferences

Dependent variable: preference for the coordination game				
	LPM		Probit	
	(1)	(2)	(3)	(4)
Own	0.049***	0.048***	0.137***	0.152***
Cooperation	(0.013)	(0.012)	(0.040)	(0.042)
			[0.048]***	[0.049]***
Partners'	−0.035***	−0.032**	−0.097**	−0.097**
Cooperation	(0.013)	(0.012)	(0.039)	(0.040)
			[−0.035]***	[−0.032]**
Logic		0.105***		0.311***
		(0.027)		(0.081)
				[0.101]***
Random		−0.007		−0.014
Dictator		(0.050)		(0.151)
				[−0.004]
Female		−0.066		−0.235
		(0.062)		(0.192)
				[−0.076]
Age		−0.011*		−0.034**
		(0.006)		(0.017)
				[−0.011]**
Economics		0.058		0.158
Student		(0.053)		(0.164)
				[0.051]
Constant	0.597***	0.717***	0.249**	0.673
	(0.045)	(0.170)	(0.121)	(0.515)
N	280	272	280	272
R^2	0.055	0.154		
Pseudo R^2			0.043	0.128

The dependent variable is equal to one if a subject chose the coordination game. Own and partners' cooperation range from 0 to 10 for each round of cooperation in stage 1. Logic ranges from 0 to 3 for each correct cognitive reflection question. Marginal effects in brackets. Subjects who did not specify their gender are excluded in (2) and (4)

Standard errors (in parentheses) clustered at group level. * $p < 0.10$ ** $p < 0.05$, *** $p < 0.01$

modification preference nor in cooperation in stage 3 between the elected representatives and the rest of the subjects ($p = 0.363$ and $p = 0.790$, respectively).

Result 2 Elected representatives are not more cooperative than other subjects.

Table 4 Cooperation Rates By Vote Outcome

Cooperation Rates in Round 11 (in %)				
	EndoMod	*EndoNot*	*ExoMod*	*ExoNot*
Individual vote				
Representative democracy (ID)				
No	64.29	23.08	40.00	46.67
	[14]	[13]	[10]	[15]
Yes	94.12	27.27	81.82	61.90
	[34]	[11]	[22]	[21]
All	85.42	25.00	68.75	55.56
	[48]	[24]	[32]	[36]
Random dictator (RD)				
No	80.00	22.22	66.67	44.44
	[10]	[9]	[12]	[18]
Yes	80.00	66.67	84.44	84.62
	[30]	[3]	[32]	[26]
All	80.00	33.33	79.55	68.18
	[40]	[12]	[44]	[44]

Number of subjects in brackets

Regarding the difference in leadership behavior between the treatments, contrary to Hypothesis 3, the randomly appointed leaders in RD cooperate on average 6.3 out of 10 times in stage 3. The elected representatives in ID cooperate 5.8 out of 10 times and the difference is not significant ($p = 0.363$). The results thus fail to confirm the findings by Drazen and Ozbay (2019). However, leaders in both treatments considerably increase their average cooperation rate between the two stages. Cooperation of leaders in stage 3 goes up by 90 percent in the RD treatment and by 130 percent in ID.

Result 3 Elected representatives are not more cooperative than randomly appointed group leaders.

The right panel of Fig. 3 shows that cooperation rates in both treatments are higher and more stable in the coordination game than in the prisoners' dilemma ($p < 0.01$, two-sided t-tests). The payoff modification is therefore effective in fostering cooperative behavior as subjects respond to the changed incentive structure of the game and mostly coordinate on the more efficient equilibrium.

But what is the effect of the endogenous modification? Table 4 shows cooperation rates and subject numbers separated by vote stage outcomes and subjects' game preferences for both treatments in the first round after the vote. The leaders' decision was considered for 31 groups (or 124 subjects) in total. Out of those, 12 leaders in ID and 10 in RD modified the payoffs to play the coordination game with their group. The remaining 9 leaders (6 in ID and 3 in RD) chose to remain with the

prisoners' dilemma. In the exogenous condition, the payoffs were modified for 8 of the 17 groups in ID and 11 of the 22 groups in RD.

Regarding cooperation rates, round 11 as shown in Table 4 provides the most conservative results because it is the first round in which the institutional change has come into effect. Over time, behavioral differences between the outcomes could self-reinforce and inflate the estimates. The lowest cooperation rates in both treatments are observed in EndoNot, where only a quarter and a third of subjects cooperate in ID and RD, respectively. ExoNot induces cooperation rates that are more than twice as high as in EndoNot in the ID treatment. This demonstrates the flip side of democracy: endogenously chosen non-cooperative institutions lead to large-scale defection. In the coordination game, cooperation is highest in EndoMod for both treatments. Subjects cooperate more if the payoffs were modified by the group leader compared to an exogenous modification, and the difference is much larger for the elected representatives (16.6 percentage points) compared to the random dictators (0.5 percentage points).

While cooperation rates in all four vote stage outcomes are significantly different from zero (see regression results in Table 5), there are not always significant differences between the outcomes, especially when controlling for modification preferences (see p-values and sharpened q-values in Table 6). Behavior under ExoNot is statistically indistinguishable from ExoMod in both treatments, even though there are two different payoff schemes at play. The difference between EndoMod and ExoMod is significant at the 10-percent level in ID. Thus, the implementation procedure through the elected representative versus the computer is of relevance to subjects. In RD, no significant effect can be found between EndoMod and ExoMod; the behavior is virtually identical in both outcomes.

However, the simple difference between cooperation rates in EndoMod and ExoMod is not an appropriate measure of the effect of the democratic policy selection. Because the leaders actively choose their games in the endogenous condition, the treatment assignment is not random. As reported in Result 1, cooperative subjects are more likely to prefer the coordination game. Therefore, the difference between the endogenous and exogenous conditions is potentially biased by self-selection. Table 4 shows that the yes-voters cooperate more than the no-voters. It is thus necessary to account for differences in group composition between the outcomes when comparing cooperation rates and deriving the effect of the endogenous policy choice.

How can the effect of the endogenous policy choice on cooperation be identified? DFP (2010) propose an identification strategy that breaks down the differences in cooperation rates between the vote stage outcomes while accounting for differences in the distribution of Yes- and No-voters. The strategy allows to estimate the increase in cooperation in response to the decision-making procedures in ID and RD that cannot be explained by the payoff change or differences in group composition. This unbiased estimate is called the democracy premium. It can be used to break down the total policy effect of endogenously changing the game to a selection effect, the exogenous treatment effect, and the democracy premium. This is done by using weighted averages of the individual cooperation rates and voter shares after the vote stage in Table 4 (the detailed calculation can be found in the Electronic

Table 5 The Effect of Democracy

Dependent variable: cooperation in round 11				
	(1)	(2)	(3)	(4)
	ID	RD	ID	RD
EndoMod	0.854***	0.800***		
	(0.063)	(0.069)		
EndoNot	0.250***	0.333***		
	(0.090)	(0.126)		
ExoMod	0.688***	0.795***		
	(0.078)	(0.066)		
ExoNot	0.556***	0.682***		
	(0.073)	(0.066)		
EndoModn			0.643***	0.800***
			(0.114)	(0.133)
EndoNotn			0.231*	0.222
			(0.118)	(0.140)
ExoModn			0.400***	0.667***
			(0.135)	(0.121)
ExoNotn			0.467***	0.444***
			(0.110)	(0.099)
EndoMody			0.941***	0.800***
			(0.073)	(0.077)
EndoNoty			0.273**	0.667***
			(0.129)	(0.243)
ExoMody			0.818***	0.844***
			(0.091)	(0.074)
ExoNoty			0.619***	0.846***
			(0.093)	(0.082)
N	140	140	140	140
R^2	0.71	0.74	0.73	0.77

OLS results, dependent variable equal to one for cooperation. Independent variables are binary indicators for vote stage outcomes. Models estimated without constant. Suffixes denote interactions with individual preferences for (-y) or against (-n) modification

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Supplementary Material). The weighting by modification preferences accounts for differences in group composition: as modification preference and cooperation are positively correlated, a group with more yes-voters can be expected to have higher cooperation rates. Therefore, the estimation approach by DFP (2010) holds voter shares constant across vote stage outcomes.

Table 7 gives an overview of the decomposed treatment effects in ID and RD. The total effect of the policy – the endogenous change from one game to another – is

Table 6 The Effect of Democracy: p-values and q-values

	(1) ID	(2) RD	(3) ID	(4) RD
p-values				
EndoNot = ExoNot	0.009***	0.015**		
EndoMod = ExoMod	0.099*	0.962		
EndoMod = EndoNot	0.000***	0.001***		
ExoMod = ExoNot	0.219	0.223		
EndoNotn = ExoNotn			0.146	0.198
EndoModn = ExoModn			0.171	0.460
EndoModn = EndoNotn			0.013**	0.003***
ExoModn = ExoNotn			0.702	0.485
EndoNoty = ExoNoty			0.031**	0.159
EndoMody = ExoMody			0.293	0.683
EndoMody = EndoNoty			0.00***	0.601
ExoMody = ExoNoty			0.128	0.983
Sharpened q-values				
EndoNot = ExoNot	0.015**	0.024**		
EndoMod = ExoMod	0.071*	0.423		
EndoMod = EndoNot	0.001***	0.006***		
ExoMod = ExoNot	0.123	0.175		
EndoNotn = ExoNotn			0.166	0.857
EndoModn = ExoModn			0.166	1.000
EndoModn = EndoNotn			0.049**	0.028**
ExoModn = ExoNotn			0.295	0.857
EndoNoty = ExoNoty			0.066*	1.000
EndoMody = ExoMody			0.265	1.000
EndoMody = EndoNoty			0.001***	1.000
ExoMody = ExoNoty			0.166	1.000

p-values and sharpened q-values (to account for multiple hypothesis testing, see Benjamini et al. (2006), Anderson (2008)) of Wald tests for differences between vote stage outcomes based on regression results reported in Table 5

given by the difference between EndoMod and EndoNot and amounts to 60 percentage points in ID and 47 in RD. The effect of endogenously switching to the coordination game is thus substantially larger for the representative democracy. Comparing the cooperation difference in the endogenous condition to the one in the exogenous condition, we find that the total policy effect is highly significant in both treatments.

The selection effect captures the higher cooperation that would be observed in the EndoNot condition if the share of yes-voters was the same as in the EndoMod groups. Under the assumption that yes-voters cooperate more, the group composition matters for the observed cooperation rates. Using the voter shares from EndoMod – where more yes-voters are present – to reweigh the average cooperation in EndoNot corrects for the self-selection of voter types into the different

Table 7 Treatment Effects – Decomposing Cooperation Rates

	Representative Democracy (ID)	Random Dictator (RD)
Total policy effect	60.42*** (10.61)	46.67*** (13.84)
Selection effect	1.05 (4.35)	22.22 (14.02)
Exogenous treatment effect	12.16 (10.51)	5.42 (5.38)
Democracy premium	47.21** (15.58)	19.07 (21.74)

Standard errors in parentheses

games. The selection effect is then the difference between the observed cooperation rates in EndoNot and the reweighted average. The difference in the proportion of player types that leads to differences in behavior is insignificant in both treatments.

The change in cooperation caused by an exogenous payoff modification is given by the exogenous treatment effect. It is given by the weighted difference between the two exogenous conditions ExoMod and ExoNot. Again, the difference is reweighted using the voter proportions of the EndoMod condition to account for potential self-selection and allow a comparison of the effects. The weighted exogenous treatment effect is again statistically insignificant in both treatments. Thus, an exogenous modification has little influence on cooperative behavior in both treatments.

Lastly, subtracting the selection effect and the exogenous treatment effect from the total policy effect, the democracy premium remains. This part of the total policy effect is the increase in cooperation not explicable by self-selection or the payoff modification itself. This residual accounts for 47 percentage points in ID and 19 in RD. The representative democracy thus induces an increase in cooperation that is more than twice that of the random dictator. This democracy premium in ID is significant and to a large extent driven by a pronounced reaction from the no-voters, who cooperate much more after endogenous modification than with the exogenous modification. The "nudge" towards cooperation from the elected group representative has a much stronger influence on behavior than modification through the computer. The more legitimate decision-making process of democratic choice seems to be especially powerful in combination with the focal figure of an elected representative. The RD democracy premium is positive as well, but not statistically significant. The democracy premium here is rather driven from the opposite direction: a significant *decrease* in cooperation in EndoNot. If anything, as opposed to the elected representative, the endogenous decision by a random dictator can have a negative impact on cooperation. Subjects do not seem to appreciate an unelected – and therefore potentially illegitimate– leader.

As a robustness check, an additional estimation approach. for the democracy premium is conducted: Bó et al. (2019) introduce an alternative identification strategy to estimate the size and significance of the effect of endogenous decision-making

Table 8 Treatment Effects–Weights-Based Analysis

Weighted average cooperation rates in round 11					
Treatment	Payoffs	Endo	Exo	Democracy effect	Standard Error (*p*-Value)
ID	Mod	83.2	68.8	14.4	8.16
		[48]	[32]		(0.078)
	Not	25.7	55.6	-29.8	8.30
		[24]	[36]		(< 0.001)
RD	Mod	80.0	79.5	0.45	6.11
		[40]	[44]		(0.941)
	Not	51.5	68.2	-17.9	7.05
		[12]	[44]		(0.011)

Number of subjects in brackets. Bootstrapped standard errors

by using weighted averages of voting behavior instead of the individual voting decisions. The cooperation rates in the endogenous conditions are reweighted according to the yes- and no-voter shares of the exogenous conditions. In both treatments the subjects in favor of the modification are overrepresented in EndoMod compared to the exogenous conditions (70.8 The difference between the weighted average cooperation rates gives the democracy effect (see Table 8). In ID, the estimated democracy effect is positive with the modified payoffs – albeit smaller than in the decomposition analysis discussed above–and significant at the 10-percent level. The democracy effect in the modified payoffs in RD is economically and statistically insignificant. There is, however, a highly significant *negative* democracy effect in both treatments: subjects are much less cooperative if the payoffs are endogenously unmodified, and the effect is almost twice as large in ID than in RD. The alternative identification strategy thus yields qualitatively the same result as the decomposition: There is a strong effect of endogenous procedures if the group leader was democratically elected.

The differences between the vote stage outcomes become even more pronounced over the course of the third stage, where behavior is self-reinforcing within groups. Figures 4 and 5 show individual cooperation rates for all four vote stage results separated by individual voting behavior and treatment over the course of the entire experiment. It can be observed from the first panel of Fig. 4 that in ID the endogenous institution has a considerable effect on those who preferred modification and leads to almost full cooperation. For both conditions the change of payoff structure results in a striking increase in willingness to cooperate. The rates are much lower without modification, especially if it was endogenously determined. The lower panel shows only individuals who were against the payoff modification in ID. The no-voters who received the coordination game through their representative drastically change their behavior and display quite stable cooperation rates. In the exogenous condition both games induce remarkably similar behavior suggesting that the democratic procedure has a stronger behavioral impact than the change in monetary incentives. Such a

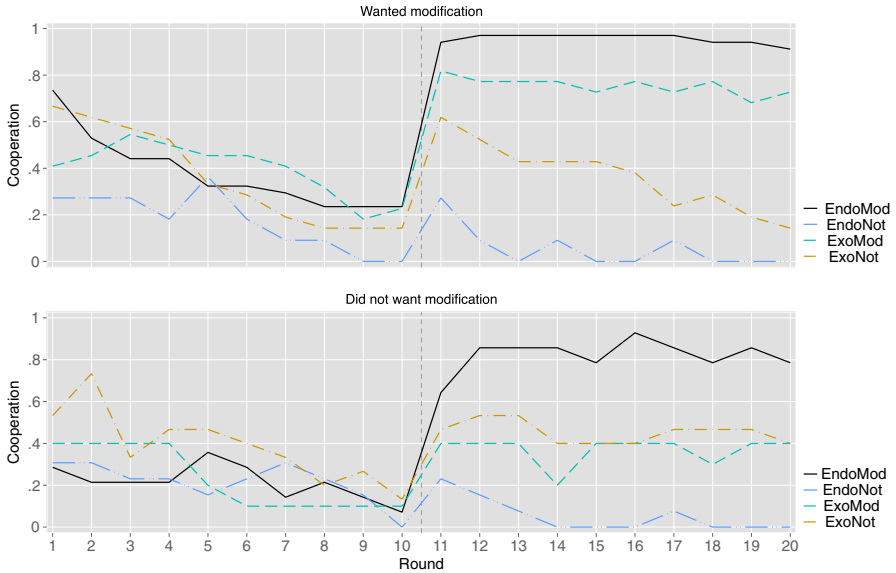

Fig. 4 Cooperation Rates – Representative Democracy (ID). *Note:* The graphs show cooperation rates for all four vote stage results separated by individual voting behavior (top panel: vote for modification, bottom panel: vote against modification)

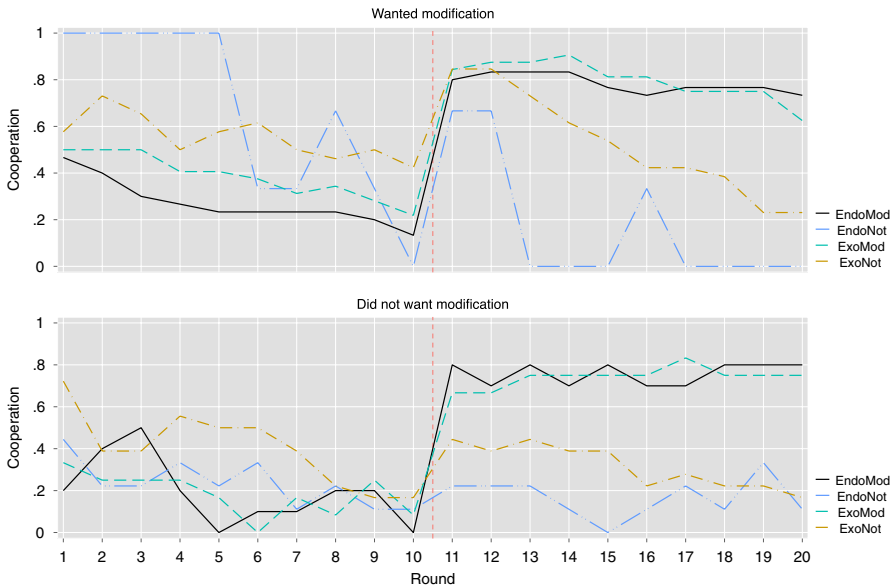

Fig. 5 Cooperation Rates – Random Dictator (RD). *Note:* The graphs show cooperation rates for all four vote stage results separated by individual voting behavior (top panel: vote for modification, bottom panel: vote against modification)

strong effect on no-voters is in line with the findings of Gallier (2020), but in contrast with DFP (2010), where the democracy premium was driven by those in favor of the policy. Figure 5 corroborates the result that the random dictator's implementation of the coordination game did not increase cooperation more than the exogenous modification: the EndoMod cooperation rates are for both voter types for a large part closely below those in ExoMod. The authority to implement a policy decision does not seem to activate compliance in the absence of democratic legitimacy. The differences between treatments are statistically significant as well: cooperation rates in EndoMod compared to ExoMod are significantly different in ID ($p < 0.001$), but not in RD ($p = 0.818$). Furthermore, cooperation in EndoMod is 14

Result 4 Cooperation is higher if the policy is introduced by a representative (ID) and the size of the democracy premium is substantial. Payoff modification by an unelected group leader (RD) does not increase cooperation.

The treatment difference between ID and RD is informative regarding the source of the democracy premium as conjectured in Hypotheses 5a and b. The evidence from the experiment is clear in this regard: the positive increase in cooperation as a response to the group leader being considered is much larger in ID. In the RD treatment, the difference between EndoMod and ExoMod is small and statistically insignificant. The legitimacy created by the election process seems to be a necessary condition for the democracy premium; authority alone cannot foster cooperation in the same way. Therefore, we reject Hypothesis 5.b and support 5.a instead: perceived legitimacy is the transmission channel of the democracy effect.

Result 5 Representative democracy is associated with a significantly larger democracy premium than randomly appointed leadership.

As an additional robustness check and to further assess the channels through which the democracy premium works, we can consider the interactions between subjects' preferences and the vote stage outcomes. One potentially relevant factor for subjects in the ID treatment is whether the elected representative was also the player that they themselves voted for. In Table 9, cooperation in the third stage is regressed on the individual modification preference, whether the individually preferred player won the election (variable *candidate*), and whether she was considered by the computer. Column (1) shows again that preferring the modification and receiving it democratically (EndoMod) both have a strongly positive influence on cooperation under the modified payoffs. It does not matter whether the preferred candidate was elected and there is no significant interaction effect between the variables. Column (2) shows that with unmodified payoffs, i.e. in the prisoners' dilemma, the modification preference has no influence on cooperation. The endogenously chosen prisoners' dilemma (EndoNot) is significantly negatively associated with cooperation. There is again neither a significant effect of having one's preferred candidate elected nor of the interaction between the modification, candidate preference and endogeneity.

Table 9 Individual Leader Support and Cooperation in ID

Dependent variable: cooperation in periods 11 to 20		
	(1)	(2)
	Modified payoffs	Unmodified payoffs
Modification	0.558***	−0.111
	(0.146)	(0.123)
Candidate	0.242	−0.092
	(0.142)	(0.091)
Modification*candidate	−0.305	−0.0100
	(0.234)	(0.087)
Endo	0.489**	−0.391**
	(0.185)	(0.178)
Modification*endo	−0.298	−0.065
	(0.195)	(0.127)
Candidate*endo	−0.0274	0.037
	(0.191)	(0.075)
Modification*candidate*endo	0.055	0.140
	(0.265)	(0.099)
Constant	0.225*	0.467**
	(0.128)	(0.175)
N	800	600
R^2	0.242	0.158

OLS results, dependent variable equal to one for cooperation. Independent variables are binary indicators for individual modification preference (*modification*), own candidate's success (*align*), and whether the representative was considered (*endo*). Standard errors (clustered at group level) in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

These results corroborate the finding that it is the decision-making procedure itself that is driving for the democracy premium.

4.3 Welfare implications

A natural next question is whether the the democracy premium corresponds to an increase in overall societal welfare. Since the exact utility functions of the subjects are unknown, we will restrict attention to monetary payoffs at first. The welfare consequences of the institutional change from one game to the other are not obvious ex-ante. The payoff modification to the coordination game, on the one hand, makes mutual cooperation more attainable and can thus increase overall payoffs. Recalling the payoffs from Table 1, mutual cooperation is the efficient outcome and gives 50 points to each player. On the other hand, if coordination fails, the penalty on unilateral defection decreases the earnings in the coordination game to 48 points as opposed to the prisoners' dilemma's deviation payoff of 60 points.

Table 10 Average Earnings per Vote Stage Outcome

Treatment	**ID**		**RD**	
Game	Prisoners' dilemma	Coordination game	Prisoners' dilemma	Coordination game
Implementation				
Endogenous	40.4	48.5	41.7	46.7
Exogenous	44.0	45.3	44.4	47.0
Total	42.6	47.2	43.8	46.9

Average earnings calculated over all ten rounds in stage 3

Comparing average earnings in stage 3, we find that in both treatments average payoffs are lower in the prisoners' dilemma (Table 10). The highest average is realized in the EndoMod condition in ID and in ExoMod in RD. However, the maximum payoff earned by one subject over the entire experiment with 53 points on average took place in the ExoMod outcome of ID. All in all, welfare is positively affected by the payoff change in both treatments. But the endogenous modification is the most efficient condition only in ID. Since average payoffs are always higher in the coordination game, it would be a natural conclusion for a social planner to circumvent the voting procedure altogether, which, after all, bears the risk of players choosing the payoff-dominated prisoners' dilemma. Instead, one could simply assign the coordination game to every group. But this comes at a cost: the highest possible cooperation rates – and therefore earnings – are only realized after the endogenous choice.

Furthermore, the monetary analysis omits some important aspects of legitimate procedures, which are at the heart of the democracy premium. For example, drawing on arguments put forward by Thibaut (1975); Sen (1995), Frey et al. (2004) introduce procedural utility to incorporate preferences about the processes that lead to instrumental outcomes into individual utility functions. The ability to exercise political participation is one source of procedural utility (Frey and Stutzer 2005). Thus, it is plausible that subjects in the experiment presented in this paper derive higher utility from having their elected representatives considered, which is something that a purely monetary welfare analysis cannot adequately capture.

5 Discussion and conclusion

The paper uses an economic experiment to quantify the influence that decision-making processes have on cooperative behavior. The effect of a decision made by an elected representative is contrasted with that of a randomly chosen group leader. Subjects are presented with the possibility of changing their payoff structure from a prisoners' dilemma into a coordination game that makes cooperation incentive-compatible. The effect of the procedure that leads to the payoffs being modified or not is given by the extent of cooperative behavior that follows the decision. The experiment uses the identification strategy developed by DFP (2010). A randomization

mechanism allows for the comparison of subjects with the same preferences, information, and incentive structure who only differ in how the incentive structure was implemented: by the group leader or by the computer. Additionally, the treatments compare the effects of elected and randomly chosen leaders.

To summarize results, the majority of subjects in both treatments prefer to modify the payoffs, and cooperative players favor the modification more, suggesting that it is appropriate to control for self-selection. In the representative democracy treatment, subjects prefer to elect pro-social representatives. Still, factually these do not behave significantly differently compared to the rest of the subjects or the randomly appointed leaders in the random dictator treatment. The findings stress the importance of procedural legitimacy over elections as selection devices: not who is the leader matters, but how the leader came into office influences behavior. Moreover, subjects cooperate more if the payoff modification is democratically introduced. In contrast to DFP (2010), the impact of the democratic policy selection is especially large for those subjects who initially did not want to introduce the modification. There is no democracy premium in the random dictator treatment. The results show that subjective legitimacy is a driving force of the democracy premium, which an unelected group leader cannot deliver.

The relevance of the results is twofold. First, the behavioral effects of institution formation are relevant for evaluating any experimental treatment effects in which subjects are assigned to different institutions. Second, the results carry policy implications. A policy that works well in one place cannot automatically be assumed to achieve similar results in another context. Further, representation is multi-faceted and how a leader is chosen determines the success of their policies to a great extent. The results from the ID treatment are reassuring in this sense: representative democracy is a widespread form of government and seems to have the largest positive impact on cooperative behavior compared to direct democracy and sortition. The sortition mechanism modeled in the RD treatment has regained popularity in recent years, mostly from grassroots movements and citizens' initiatives, especially for environmental policy questions (Dryzek and Tucker 2008; Lorent 2019; Zimmer 2021). Even though it has theoretical advantages, i.e., little proneness to corruption, the experimental results presented here imply that one should be cautious regarding the procedural legitimacy and thus the effectiveness of these mechanisms.

Can the democracy premium be explained by economic theory? Markussen et al. (2014) claim that it is rationalizable with the model of inequality aversion by Fehr and Schmidt (1999). Voting is a credible signal of an intention to cooperate that prompts inequality-averse subjects to cooperate in the coordination game (Markussen et al. 2014, p.307). However, the argument has no bite in the experiment presented in this paper because subjects are informed about their representative's intention to modify payoffs even when it is not considered. The signaling component does therefore neither differ between the endogenous and exogenous conditions nor between the two treatments. There is no reason why an elected representative should deliver stronger cues towards the cooperative equilibrium than the randomly appointed leader if we restrict attention to inequality aversion. Even when the players are assumed to be not purely maximizing their own payoffs, the procedure itself is not sufficiently consequential to account for the democracy premium. If preferences about political participation enter

the utility functions, e.g., in the form of procedural utility Frey et al. (2004) this would create a level effect on subjects' utility in the endogenous conditions. But to create the democracy premium, the procedural preferences have to interact with the treatment conditions in a way that creates differences in *behavior*, not utility. A "warm glow" feeling of political participation has to induce players to cooperate if they were considered but make them defect if they were not.

Dannenberg and Gallier (2020) suggest that the endogenous and collective implementation of an institution may evoke feelings of group identity, which can be a powerful activator of social-preferences (Akerlof and Kranton, 2000; Chen and Li, 2009). In a group-contingent social preference model, group identity can influence equilibrium selection in coordination games (Chen and Chen, 2011). Such a model could serve to explain the democracy premium under the assumption that the interaction between subjects in the first stages is not sufficient to induce group identity in the exogenous conditions, and moreover, that endogenously refusing the payoff change has adverse effects on group identity.

The perhaps most promising avenue for further theoretical research on the democracy premium is Psychological Game Theory (PGT) (Geanakoplos et al., 1989; Battigalli and Dufwenberg, 2009). PGT formally incorporates belief-dependent motivations into game theory. One especially noteworthy application is guilt aversion (Battigalli and Dufwenberg, 2007). In a guilt aversion framework, disutility is created from a failure to live up to others' expectations. The model fits well with the democracy premium if the choice to modify the payoffs is a statement of intent to cooperate. Under endogenous modification, a deviation from mutual cooperation would then be seen as "letting the other player down". It is up to further research should develop a theoretical synthesis of the numerous experimental studies on the democracy premium and their ambiguous findings. One important inconsistency in the literature on the democracy premium is who is affected most by the the democratic process: the ones in favor or the ones opposing the institutional change. Theoretical explanations hinge on this as much as policy implications derived from the experimental insights.

Limitations of the abstract experimental design presented here are that many essential features of representative democracies are excluded in the ID treatment. There is no running for elections, neither pandering nor accountability, and no rent for the elected politician. As such, the external validity of the results can be further improved. Incorporating the randomization mechanism into more complex experiments promises to deliver unbiased estimates of various kinds of endogenous treatment effects. Possible extensions to the study are giving more power to the representative, e.g., by letting her decide on the strategies of the citizens. Repeated elections that create accountability, campaigning of candidates, and preference heterogeneity are further relevant factors in representative democracies to potentially include.

Funding Open Access funding enabled and organized by Projekt DEAL.

References

Akerlof, G. A., & Kranton, R. E. (2000). Economics and Identity. *The quarterly journal of economics*, *115*(3), 715–753.

Alizada, N., Cole, R., Gastaldi L., Grahn, S., Hellmeier, S., Kolvani P., Lachapelle, J., Lührmann, A., Maerz, S. F., Pillai, S., & Lindberg, S. I. (2021). Autocratization turns viral. Democracy report 2021.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American statistical association*, *103*(484), 1481–1495.

Andreoni, J., & Gee, L. K. (2012). Gun for hire: Delegated enforcement and peer punishment in public goods provision. *Journal of Public Economics*, *96*(11–12), 1036–1046.

Bardhan, P. (2000). Irrigation and cooperation: An empirical analysis of 48 irrigation communities in South India. *Economic Development and Cultural Change*, *48*(4), 847–865.

Barnes, J. (1984). *Complete works of Aristotle*, Volume 1: The revised Oxford translation. Vol. 1. Princeton University Press.

Battigalli, P., & Dufwenberg, M. (2007). Guilt in games. *American Economic Review*, *97*(2), 170–176.

Battigalli, P., & Dufwenberg, M. (2009). Dynamic psychological games. *Journal of Economic Theory*, *144*(1), 1–35.

Benjamini, Y., Krieger A. M., & Yekutieli D., (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, *93*(3), 491-507.

Black, S. E., & Lynch, L. M., (2001). How to compete: The impact of workplace practices and information technology on productivity. *Review of Economics and Statistics*, *83*(3), 434–445.

Dal Bó, P. (2014). Experimental evidence on the workings of democratic institutions. In S. Galiani & I. Sened (Eds.), *Institutions, property rights, and economic growth: The legacy of Douglass North* (pp. 266–288). Cambridge: Cambridge University Press.

Bock, O., Ingmar B., & Nicklisch A. (2014). hroot: Hamburg registration and organization online tool. *European Economic Review*, *71*(1), 117–120.

Bonin, J. P., Derek C. Jones D. C., & Putterman L. (1993). Theoretical and empirical studies of producer cooperatives: Will ever the twain meet? *Journal of Economic Literature*, *31*(3), 1290–1320.

Dal Bó, Pedro, A. F., & Kamei, K. (2019). *The democracy effect: A weights based identification strategy*. National Bureau of Economic Research: Tech. rep.

Brandts, J., & Charness, G. (2000). Hot vs. cold: Sequential responses and preference stability in experimental games. *Experimental Economics*, *2*(3), 227–238.

Brandts, J., Cooper D. J., & Weber R. A. (2015). Legitimacy, communication, and leadership in the turnaround game. *Management Science*, *61*(11), 2627–2645.

Brandts, J., Güth, W., & Stiehler, A. (2006). I want YOU! An experiment studying motivational effects when assigning distributive power. *Labour Economics*, *13*(1), 1–17.

Brosig, J., Weimann, J., & Yang C.-L. (2003). The hot versus cold effect in a simple bargaining experiment. *Experimental Economics 6*(1), 75–90.

Carpenter, J., Graham, M., & Jesse Wolf. (2013). Cognitive ability and strategic sophistication. *Games and Economic Behavior, 80*, 115–130.

Castillo, J. G., Xu Z., Ping Z., & Xu X. (2017). Institutional legitimacy and public goods games: A laboratory experiment on the distribution of sanctioning power. *Mimeo*.

Cavalcanti, C., Schläpfer, F., & Schmid B. (2010). Public participation and willingness to cooperate in common-pool resource management: A field Experiment with fishing communities in Brazil. *Ecological Economics*, *69*(3), 613–622.

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. *Experimental Economics*, *14*(1), 47–83.

Chen, R., & Chen, Y. (2011). The potential of social identity for equilibrium selection. *American Economic Review*, *101*(6), 2562–2589.

Chen, Y., & Li, SX. (2009). Group identity and social preferences. *American Economic Review*, *99*(1), 431–457.
Cooper, R., DeJong D. V., Forsythe R., & Ross T. W. (1996). Cooperation without reputation: Experimental evidence from prisoner's dilemma games. *Games and Economic Behavior*, *12*(2), 187–218.
Corazzini, L., Kube S., André Maréchal M., & Nicolo A, (2014). Elections and deceptions: An experimental study on the behavioral effects of democracy. *American Journal of Political Science*, *58*(3), 579–592.
Dal Bó, P., Foster, A., & Putterman, L. (2010). Institutions and behavior: Experimental evidence on the effects of democracy. *American Economic Review*, *100*(5), 2205–2229.
Dannenberg, A., & Gallier, C. (2020). The choice of institutions to solve cooperation problems: A survey of experimental research. *Experimental Economics, 23,* 716–749.
Dannenberg, A., Haita-Falah, C., & Zitzelsberger, S. (2020). Voting on the threat of exclusion in a public goods experiment. *Experimental Economics, 23,* 84–109.
Drazen, A., & Ozbay, E. Y. (2019). Does being chosen to lead induce non-selfish behavior? Experimental evidence on reciprocity. *Journal of Public Economics, 174,* 13–21.
Dryzek, J. S., & Tucker, A. (2008). Deliberative innovation to different effect: Consensus conferences in Denmark, France, and the United States. *Public Administration Review*, *68*(5), 864–876.
Fearon, J. D., Humphreys, M., & Weinstein J. M. (2011). Democratic institutions and collective action capacity: Results from a field experiment in postconflict liberia. *APSA 2011 Annual meeting paper.*
Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, *114*(3), 817–868.
Fischbacher, U., S. Gächter, & Quercia S. (2012). The behavioral validity of the strategy method in public good experiments. *Journal of Economic Psychology*, *33*(4), 897–913.
Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42.
Frey, B. S., & Stutzer, A. (2005). Beyond outcomes: Measuring procedural utility. *Oxford Economic Papers*, *57*(1), 90–111.
Frey, B. S., Benz M., & Stutzer A. (2004). Introducing procedural utility: Not only what, but also how matters. *Journal of Institutional and Theoretical Economics*, *160*(3), 377–401.
Gallier, C., (2020). Democracy and compliance in public goods games. *European Economic Review*, *121*, 103346.
Geanakoplos, J., Pearce D., & Stacchetti E. (1989). Psychological games and sequential rationality. *Games and Economic Behavior*, *1*(1), 60–79.
Grossman, G., & Baldassarri, D. (2012). The impact of elections on cooperation: Evidence from a Lab-in-the-field experiment in uganda. *American Journal of Political Science*, *56*(4), 964–985.
Hamman, J. R., Weber R. A., & Woon J. (2011). An experimental investigation of electoral delegation and the provision of public goods. *American Journal of Political Science*, *55*(4), 738–752.
Kamei, K. (2017). Collective selection of representatives and the democracy effect. *Mimeo*.
Kamei, K. (2016). Democracy and resilient pro-social behavioral change: An experimental study. *Social Choice and Welfare*, *47*(2), 359–378.
Kamei, K., Putterman, L., & Tyran J-R (2015). State or nature? Endogenous formal versus informal sanctions in the voluntary provision of public goods. *Experimental Economics*, *18*(1), 38–65.
Locke, J. (1983). *The second treatise of government*. Stuttgart: Reclam.
Marcin, I., Robalo, P., & Tausch, F. (2019). Institutional endogeneity and third-party punishment in social dilemmas. *Journal of Economic Behavior & Organization, 161,* 243–264.
Markussen, T., Putterman, L., & Tyran J-R. (2014). Self-organization for collective action: An experimental study of voting on sanction regimes. *Review of Economic Studies*, *81*(1), 301–324.
Mechtenberg, L., & Jean-Robert, T. (2019). Voter Motivation and the quality of democratic choice. *Games and Economic Behavior*, *116*.
Olken, B. A. (2010). Direct democracy and local public goods: Evidence from a field experiment in Indonesia. *American Political Science Review*, *104*(2), 243–267.
Ostrom, E., Walker, J., & Roy Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. *American Political Science Review*, *86*(2), 404–417.
Ostrom, E. (1991). *Governing the commons*. Cambridge: Cambridge University Press.
Pascal L. (2019). La Communauté germanophone se dote d'une assemblée citoyenne. *Le Soir*. Accessed: 2021-04-11.
Rivas, M. F., & Sutter, M. (2011). The benefits of voluntary leadership in experimental public goods games. *Economics Letters*, *112*(2), 176–178.

Schmidt, V. A. (2013). Democracy and legitimacy in the european union revisited: Input, output and throughput. *Political Studies*, *61*(1), 2–22.
Sen, A. (1995). Rationality and social choice. *American Economic Review*, *85*(1), 1–24.
Sutter, M., Haigner, S., & Kocher, M. G. (2010). Choosing the carrot or the stick? Endogenous institutional choice in social dilemma situations. *The Review of Economic Studies*, *77*(4), 1540–1566.
Thibaut, J. W. L. W. (1975). *Procedural justice: A psychological analysis*. Hillsdale: Lawrence Erlbaum Associates.
Towfigh, E. V., Goerg S. J., Andreas Glöckner, P. L., Llorente-Saguer, A., Bade, S. & Kurschilgen C. (2016). Do direct-democratic procedures lead to higher acceptance than political representation?" *Public Choice*, *167*(1-2), 47-65.
Tyran, J.R., & Feld, L. P. (2006). Achieving compliance when legal sanctions are non-deterrent. *The Scandinavian Journal of Economics*, *108*(1), 135–156.
van Ham C., Thomassen, J. J., Aarts K, & Andeweg R. B. (2017). *Myth and reality of the legitimacy crisis: Explaining trends and cross-national differences in established democracies*. Oxford University Press.
Vollan, B., Landmann, A., Zhou, Y., Biliang, H., & Herrmann-Pillath, C. (2017). Cooperation and authoritarian values: An experimental study in China. *European Economic Review, 93,* 90–105.
Zimmer, S. (2021). Klimaneustart Berlin. https://klimaneustart.berlin/klimabuergerinnenrat/. Accessed: 2021-04-11.