



Public good bargaining under mandatory and discretionary rules: experimental evidence

John Duffy^{1,2} · SunTak Kim³ 

Received: 17 January 2022 / Revised: 3 November 2023 / Accepted: 16 November 2023 /
Published online: 12 December 2023

© The Author(s), under exclusive licence to Economic Science Association 2023

Abstract

We experimentally test a model of public good bargaining due to Bowen et al. (Am Econ Rev 104:2941–2974, 2014) and compare two institutions governing bargaining over public good allocations. The setup involves two parties negotiating the distribution of a fixed endowment between a public good and each party's individual account. Parties attach either high or low weight to the public good and the difference in these weights reflects the degree of polarization. Under discretionary bargaining rules, the status quo default allocation to the group account (in the event of disagreement) is zero while under the mandatory bargaining rule it is equal to the level last agreed upon. The mandatory rule thus creates a dynamic relationship between current decisions and future payoffs, and our experiment tests the theoretical prediction that the efficient level of public good is provided under the mandatory rule while the level of public good funding is at a sub-optimal level under the discretionary rule. Consistent with the theory, we find that proposers (particularly those attaching high weight to the public good) propose significantly greater allocations to the public good under mandatory rules than under discretionary rules and this result is strengthened with an increase in polarization. Still, public good allocations under mandatory rules fall short of steady state predictions, primarily due to fairness concerns that prevent proposers from exercising full proposer power.

Keywords Public goods · Dynamic bargaining game · Mandatory rules · Discretionary rules · Bargaining experiment

JEL Classification C78 · C92 · E62 · H41 · H61

The replication material for this project is available at: <https://osf.io/3su4p/> Funding for this project was provided by the UC Irvine School of Social Sciences. The experimental protocol was approved by the UC Irvine Institutional Review Board, HS#2011-8378. Jason Ralston provided expert research assistance. We thank the Editor, Andrzej Baranski-Madrigal, and two referees for helpful comments and suggestions. Special thanks go to Renee Bowen, Ying Chen and Hülya Eraslan for sharing their Matlab code to generate Markov perfect equilibria.

Extended author information available on the last page of the article

1 Introduction

Expenditures on public goods are often the result of a bargaining process between legislative parties that differ in terms of their bargaining power and in the utility they derive from public good expenditures. In addition, there is often a dynamic element to such negotiations in that a considerable fraction of public expenditures are often *mandated* by law, and are not subject to discretionary, renegotiation from one period to the next. For example, in the U.S., mandatory expenditures (e.g., on Social Security, Medicare, Medicaid, Veteran's Benefits and other sources of "Income Security" - the latter including Disability Assistance, Food and Nutrition Assistance, Supplemental Security Income, Earned Income Tax Credits, Child Tax Credits, Unemployment Insurance, Student Loans, and Deposit Insurance) accounted for over two thirds of total federal government spending in FY 2020, while discretionary spending that is subject to renegotiation each year (e.g., on military and non-defense cabinet offices) accounted for about one quarter of FY 2020 spending—see Fig. 1 (The remaining amount is interest on government debt which is not included in the figure).

According to the Congressional Budget Office¹ U.S. federal mandatory spending has steadily increased over time from an average of 12.7 percent of GDP over the years 2010–2019 to more than 20% of GDP during the pandemic years of 2020–21 and are projected to be 15.2 percent of GDP by 2030. By contrast, discretionary outlays have steadily declined from 7 percent of GDP in 2013 to 6.4 percent in 2020 and are projected to be 5.6 percent of GDP in 2030.

In this paper we explore the process by which two parties bargain over public good expenditures under two distinct budgeting rules. Specifically, we experimentally test a model of this bargaining process due to Bowen et al. (2014), henceforth "BCE". Under a purely discretionary bargaining rule, BCE assume that the status quo allocation to the public good in the event of a bargaining disagreement is always zero. However, under mandatory bargaining rules, the status quo default public

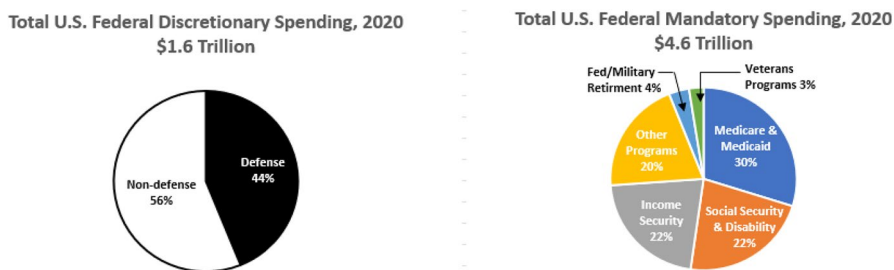


Fig. 1 Breakdown of U.S. Federal Government Spending in 2020 *Source:* Congressional Budget Office Data

¹ CBO, The Budget and Economic Outlook: 2022 to 2032, May 2022, <https://www.cbo.gov/publication/57950>.

good expenditure in the event that there is no bargaining agreement is assumed to be equal to the level of public good expenditure that was last agreed upon by the two parties. BCE show that under this mandatory bargaining rule, allocations to the public good are higher and can Pareto dominate allocations under the discretionary rule under certain conditions. BCE thus provide a simple dynamic mechanism that enables efficient provision of public goods to be attained and rationalizes the steady growth of mandatory spending in the historical U.S. federal budget. Our aim in this paper is to experimentally test the predictions of the BCE model in a laboratory experiment with paid human subjects. We implement a version of their model in the laboratory and we find strong, though imperfect support for the model predictions in our experimental data.

Our paper is most closely related to the experimental literature on coalitional and legislative bargaining; see, e.g. Palfrey (2015) and Baranski and Morton (2022) for surveys. John Kagel has made many pioneering contributions to this literature including Fréchette et al. (2003, 2005a, b, c, 2012) and Baranski and Kagel (2015). Of these papers, the one that is most closely related to this paper is Fréchette et al. (2012). They consider a version of the Baron and Ferejohn (1989) model of majoritarian coalitional bargaining where 5 players must make and vote on allocations to both private (particularistic) and public goods as in our study. The difference between their paper and ours is that we consider dynamic budgeting rules that depend on the status quo level of previously agreed upon public good expenditures and we have only 2 players who make or agree to proposed allocations, so that our decision rule amounts to unanimity. Further, in our setting, following BCE, the public good yields a nonlinear payoff (implying an interior optimum for the public good amount) that varies with the player type—the high (low) type gets a higher (lower) utility from the public good.

Our paper is also related to the recent literature on dynamic bargaining experiments, possibly with an endogenous status quo, in legislative and/or multilateral settings (e.g., Battaglini et al. (2012) and Battaglini and Palfrey (2012)). Those papers allow for an endogenous status quo, but focus either on purely distributive politics without a public good element or on the provision of durable public goods under different voting rules (majority/unanimity).

The main question we address in this paper is whether dynamic, mandatory budgeting rules matter for the achievement of the efficient level of the public good relative to discretionary budgeting rules when the public good allocation is the result of a dynamic bargaining process by two parties with different interests.

Our experimental data clearly show that Pareto improvements in public good allocations are possible under dynamic, mandatory budget rules, as opposed to discretionary rules. These improvements result from private negotiations between interested parties and occur in the absence of transaction costs. In this sense, the dynamic public good bargaining game provides a mechanism to obtain efficient public good provision in line with the Coase Theorem. To preview our results, we find that in the discretionary treatment, participants tend to allocate more to the public good than what is anticipated based on the static equilibrium. However, in the mandatory treatments, participants allocate even more to the public good, and come very close to achieving the Pareto efficient outcome in public good provision. However,

they fall just short of that level, and we attribute this failure to fairness concerns. In order to convince responders to accept proposals, proposers are not able to exercise full proposer power. Instead, proposers must award responders with some private points, despite the equilibrium prediction that proposers allocate zero private points to responders in most cases unless the responders have a strong outside option. These private points awarded to responders reduce both the public good allocation and the proposer's own private allocations in the mandatory treatments, and that is why the Pareto optimum is not quite achieved.

Our results provide some support for the key insight of BCE that the endogenous status quo level of public good provision works as an outside option for responders in bargaining under the mandatory budget rules. Proposers have an incentive to maintain or increase allocations to the public good over the current status quo as insurance against the possibility that they lose their status as a proposer in the future (the roles of proposer and responder change with a fixed probability in each round of our dynamic bargaining supergames). This is why public good provision can grow close to the Pareto efficient level in the steady states of dynamic bargaining games under the mandatory budget rules.

2 Model and experimental design

The model we implement in the laboratory was originally proposed as a dynamic game of public good bargaining by Bowen et al. (2014) (BCE). It involves bargaining between two parties about the allocation of an endowment across both public and private accounts under alternative budget rules and over an indefinite horizon.

Specifically, two parties repeatedly bargain with one another in an indefinite sequence of rounds over how to allocate a fixed endowment—in our experiment 100 points in each round—across a group account (public good) and two private accounts, one for each of the two parties. The points assigned to the group account contribute to the earnings of both members of the pair, while the points assigned to each of the two private accounts only accrue to the earnings of the individual parties who own those private accounts.

At the start of each new sequence (supergame) of rounds, the two members of each party are randomly paired and are equally likely to be chosen to be the proposer (the other player is the responder). Following the first round of the sequence, if the game continues, the current proposer continues to be the proposer in the next round with probability p , and with probability $1 - p$, the proposer and responder switch roles. We chose to set $p = .60$ throughout all treatments of the experiment so that there is some persistence to players remaining in the same proposer/responder roles from one round to the next of each sequence, but also allowing for political change (i.e., changes in the majority party which monopolizes the proposer power) to occur, here with probability $1 - p = .40$.

The players in each pair are also randomly assigned to be either a high or a low type player, which refers to how they value the public good (as explained below). Each pair has one high and one low type player and this designation does not change over the course of the supergame.

In each pair, the proposer chooses an allocation of the 100 points (endowed anew each round) to the “group” account, the “private” account of the proposer and the private account of the responder in each round. The allocations across all three accounts must sum to exactly 100 points. If the responder (matched with the proposer) accepts this proposal, the round payoffs for the proposer and responder in the pair are realized according to the agreed upon proposal. A player with X points in their own private account and Y points in the group account would have round earnings calculated as follows:

$$\text{Player points earned} = X + \theta_i \ln Y, \quad i \in \{L, H\}.$$

Here, if the player is a “low” type, then $\theta_L = 25$ in all treatments. If the player is a “high” type, then $\theta_H = 40$ or $\theta_H = 55$ depending on the treatment conditions. $\ln$ refers to the natural logarithm.

Once a sequence ended, subjects were randomly rematched into new pairs and their types (high or low) and initial assignments as either the proposer or responder were newly and randomly assigned at the start of the next sequence. In this way, subjects in each experimental session played multiple indefinitely repeated games (supergames) of public and private good bargaining as either high or low types (in terms of their valuation for the public good) and also traded off proposer and responder roles according to the Markovian switching probability p as described above.

Our experiment consists of two treatment variables, (1) the *budget rule*, which is either “discretionary” or “mandatory” as discussed in further detail below and (2) for the mandatory treatments only, the degree of political polarization which is measured by the difference, $\theta_H - \theta_L$. As noted earlier, we always have $\theta_L = 25$. In the baseline “aligned” treatment the high type has $\theta_H = 40$ and in the “polarized” treatment, the high type has $\theta_H = 55$ (accordingly $\theta_H - \theta_L$ is larger in the polarized treatment). Note that the discretionary treatment uses the same θ values as the baseline aligned treatment, $(\theta_H, \theta_L) = (40, 25)$.

At the start of each new sequence, the default number of points, Y , in the group account is set to one in both the mandatory and discretionary treatments. Thus, under the logarithmic specification for the public good component for the stage utility function, there will be a zero public good payoff from this default level. The default number of points in the two private accounts (which enter utility linearly) are both 0.

Following BCE, we distinguish alternative budget rules from one other according to whether the public good levels that are agreed upon in previous rounds of a given sequence are persistent or not. If there is disagreement under the *discretionary* budget rules, then the points allocated to the group account are reset to one and both private accounts are reset to zero. In other words, there is no persistence to public good levels in the discretionary treatment.

By contrast, under the *mandatory* budget rules, if a proposal is accepted, the accepted amount in the group account becomes the new status quo default public good amount for future rounds of that same sequence. In the event that there is disagreement about future proposals within that same sequence (supergame) then,

under the mandatory rule, both the proposer and the responder's private accounts default to zero points but the group account defaults to the status quo level - the most recently agreed to public good allocation in the sequence—so that both players may still receive some positive utility benefit from a disagreement outcome.

Thus, we implement three treatments. In the discretionary, aligned (D) treatment (just called the “discretionary treatment” hereafter), $(\theta_H, \theta_L) = (40, 25)$ and the discretionary budgeting rule is in place. In the mandatory-aligned treatment (Ma) $(\theta_H, \theta_L) = (40, 25)$ and in the mandatory-polarized treatment (Mp) $(\theta_H, \theta_L) = (55, 25)$, and in these two treatments, the mandatory budgeting rule is in place.² These parameter choices imply that the Pareto efficient level for the public good allocation is $\theta_H + \theta_L = 65$ in the aligned treatment and 80 in the polarized treatment.³ We required the proposer to allocate at least 1 point to the group account to prevent outcomes with negative payoffs ($\ln Y$ for $0 \leq Y < 1$ can result in a large negative number). Note one difference of this theory from standard public good or voluntary contribution games is that the utility from the public good allocation is nonlinear (logarithmic) which enables unique interior solutions; utility from private point allocations is linear as is more typical in those games.⁴ Finally, our design is between-subjects; each session consisted of 10 subjects who participated in multiple supergames or “sequences” all conducted under the same treatment conditions (i.e., the budget rule, discretionary or mandatory, and the values of the public good weighting parameters (θ_H, θ_L) are held fixed in every session of the treatments).

At the end of each round of a sequence there is a one-fifth chance that the current sequence does not continue on with another round. We thus implement bargaining over an indefinite horizon with a discount factor of $\delta = .80$ using the method of random termination. After learning whether the most recent proposal was accepted or not, subjects were shown a randomly drawn integer from 1-5 inclusive at the end of each round. They were instructed that if a 5 was drawn then the sequence would end; otherwise the sequence would continue with another round and in that case, the status quo level for the public good in the mandatory treatments would carry forward as well.

We drew the random numbers in advance and we used several different sequences of random number draws across both the discretionary and mandatory treatment sessions. This design ensures that the length of sequences are the same between mandatory and a discretionary sessions so that we can more readily compare the dynamic

² Our mandatory-aligned and mandatory-polarized treatments correspond to the ‘low-polarization’ and ‘high-polarization’ cases of mandatory budget rules in Bowen et al. (2014).

³ Under the induced logarithmic utility specification for the public good payoff, it can be easily shown that the Pareto efficient level for the public good allocation in the dynamic bargaining games described in this section is $\theta_H + \theta_L$; see Bowen et al., (2014) Sec.II.

⁴ The experimental literature exploring Baron-Ferejohn type legislative bargaining usually involves allocation of private points *only* among three or more players. As noted earlier, Fréchet et al. (2012) is an exception in that they allow both public and private (particularistic) goods, similar to our study, but in a multilateral bargaining setting. While they examine static multi-stage bargaining games, their utility is linear in both the private and public goods.

Table 1 Number of realized rounds

D1,2, Ma1,2, Mp1,2		D3,4, Ma3,4, Mp3,4		D5, Ma5, Mp5	
Sequence	Rounds	Sequence	Rounds	Sequence	Rounds
1	3	1	6	1	3
2	4	2	1	2	8
3	4	3	5	3	6
4	6	4	2	4	5
5	12	5	10	5	1
6	6	6	6	6	10
7	4	7	3	7	3
		8	3	8	3
		9	5		
	39		41		39

data between the different treatments. The realized number of rounds for our sessions are as shown in Table 1.

For instance, for the first two sessions, 1-2 of treatments D, Ma, and Mp, we had 7 sequences (supergames) lasting various numbers of rounds that summed to 39 rounds in total.

At the end of a session, we randomly chose two sequences from all sequences played in a session and we paid subjects according to the points they earned in the final rounds of the two chosen sequences.⁵ The points subjects earned in those two final rounds were converted into money at the fixed and known rate of 15 points = US\$1 and the point totals thus calculated were paid together with a \$7 show-up fee.

The experiment was computerized and programmed using oTree (Chen et al., 2016). On the relevant decision screens, we reminded subjects of the history of all group (public good) and private points in the previous rounds as well as the status quo public good levels to aid them in making decisions. They also had access to online calculators.

All sessions were conducted in the Experimental Social Science Laboratory (ESSL) at UC Irvine. Prior to making any decisions on the networked computer workstations of the laboratory, subjects were given written instructions which were also read aloud. See Appendix A for a copy of these instructions for the aligned treatment (both the Discretionary and Mandatory versions). After the instructions were read, subjects completed a quiz (which can be found at the end of the instructions given in Appendix A). Subjects' quiz answers were reviewed by the experimenter; if a subject got a quiz question wrong, the experimenter went over the correct answer with the subject before the experiment began.

Subjects were undergraduate students at UC Irvine pursuing a variety of different major programs of study. They were recruited using the Sona systems software.

⁵ Here we follow the practice recommended by Sherstyuk et al. (2013).

Each subject participated in just one session. Total average earnings (including the show-up payment) were \$24.03 for a two-hour experimental session.

3 Equilibrium and hypotheses

In the dynamic bargaining game described in Sect. 2, subjects should maximize their discounted payoffs (discount factor $\delta = .80$) over an indefinite sequence of allocation decisions with the induced stage utility given by $u_t = X_t + \theta_i \ln Y_t$. Assuming they do so, Fig. 2 shows the resulting Markov perfect equilibrium public good allocations. These equilibrium allocations are plotted as a function of the status quo level for the public good (on the horizontal axis) for the two mandatory treatments (aligned, Ma and polarized, Mp) for the parameter (θ) values that we used in the experiment. Figure 2 shows the predicted public good allocations, Y , proposed by both high and low type proposers. For most status quo levels, the proposer should allocate the remainder of the endowment to his own private account, giving zero to the responder, thereby exercising *full proposer power*. Figure B.1 in the Appendix presents predicted private point allocations X for both proposers and responders in the two mandatory treatments.

These are of course, the rational actor model predictions under standard, money maximizing preferences as specified above. In Appendix C, we show that a modified version of the discretionary model with other-regarding preferences results in a slight increase in public good allocations and less than full exercise of proposer power.

In the equilibrium for the discretionary (aligned) treatment (which is not shown in any figure), each type of player proposes his static equilibrium level for the public good, $Y^* = \theta_i$. The logic here directly follows from the first order condition from the static, one-shot maximization problem which yields $-1 + \frac{\theta_i}{Y} = 0$, or $Y = \theta_i$. Thus, in the discretionary aligned treatment $Y^* = 40$ or 25 depending on whether the proposer is a high or low type (the interior optimum for the stage utility with the assumption of full proposer power $X_{t,proposer} = 100 - Y_t$), and all remaining points go to the proposer's own private account (See Proposition 1 of Bowen et al. (2014)).

Under the mandatory rule, the two parties have an incentive to maintain the level of public good provision at least as high as the status quo level (the so-called *status quo effect*), as illustrated in Fig. 2. Maintenance of the current public good allocations provides proposers with some insurance against the future possibility of losing proposer power (which happens with probability $1 - p$ in the next round if one was a proposer in the current round). Once in the role of a responder, a higher status quo level for public goods reinforces the player's bargaining power, and anticipating this, proposers operating under the mandatory rule have an incentive to push up or maintain public good allocations, relative to those under the discretionary rule.

This dynamic incentive results in the eventual growth of public good amounts up to the Pareto efficient level ($\theta_H + \theta_L$) under the mandatory regimes. The achievement of the Pareto efficient level of public good provision stands in contrast to the

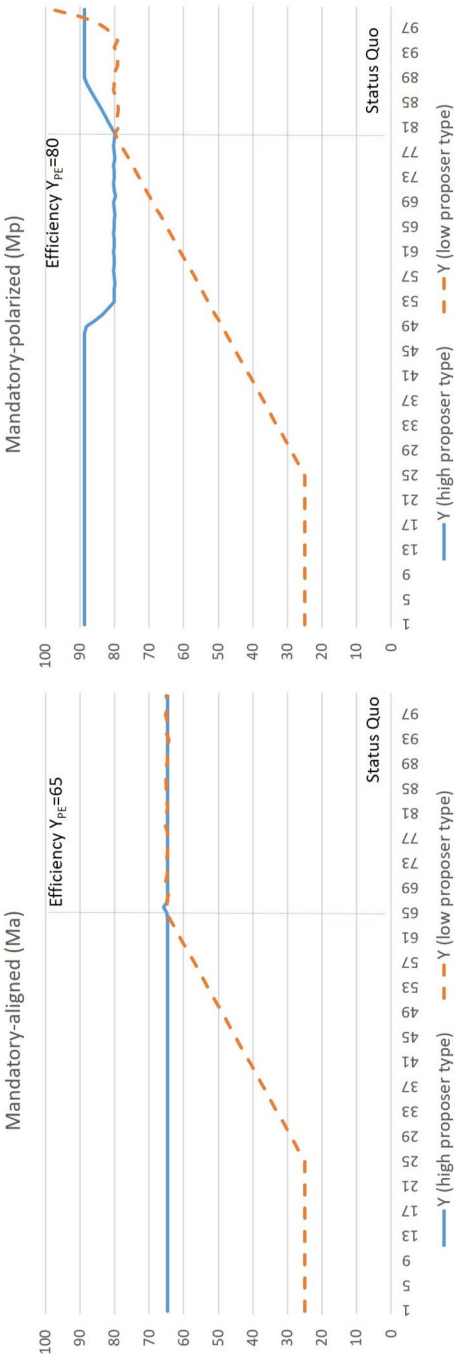


Fig. 2 Markov perfect equilibrium predictions for public good allocations in the two mandatory treatments, as a function of the status quo level. left panel: mandatory-aligned (ma); right panel: mandatory-polarized (Mp)

perpetual oscillation between each type's static equilibrium levels of public good provision (θ_H and θ_L) that is predicted to occur under the discretionary budget rule.

To clarify these differences, Bowen et al. (2014) introduce the notion of the "dynamic optimum" (Y_{DO}^*) which is roughly the public good allocations that maximize the dynamic payoff of each proposer type (*high* or *low*) under full proposer power;

$$Y_{DO}^*(high) = \frac{1 + \delta - 2\delta p}{1 - \delta p} \theta_H, \quad Y_{DO}^*(low) = \theta_L$$

where δ is discount factor, or probability of random termination in our experiment, and p is the Markov probability of the roles switching between proposer and responder from round to round. It is always the case that $Y_{DO}^*(high) > \theta_H$ unless $\delta = p = 0$.⁶

Basically, if the status quo level of the public good is below the (type-specific) dynamic optimum, each type has an incentive to raise public goods to their own dynamic optimum level immediately; if the status quo level is above the dynamic optimum but below the Pareto efficient level, then each type will maintain the current status quo; finally, for a status quo level above Pareto efficiency, both types propose the efficient level in equilibrium (see Fig. 2). These patterns for equilibrium public good offers largely hold without exception in the mandatory-aligned (Ma) case. However, the high type's equilibrium public good proposals may *overshoot* the Pareto efficient level when the status quo is lower (below half of the endowment) or is above the efficient level in the mandatory-polarized (Mp) case (and there is a small region of irregularity in low type's proposals for status quo levles close to the full endowment of 100). With the initial status quo being 1 point in the group account at the start of each sequence (which is the case for all treatments of our experiment), the equilibrium dynamics predict that the public good allocations in the steady states are equal to the high type's dynamic optimum: $Y_{SS}^* = Y_{DO}^*(high) \approx 64.615$ in the Ma treatment and the Pareto efficient level $Y_{SS}^* = \theta_H + \theta_L = 80$ in the Mp treatment, respectively, for our parameter choices. Markov perfect equilibrium is formally characterized in Proposition 3 (Ma or *low-polarization*) and Proposition 4 (Mp or *high-polarization*), and equilibrium steady states are characterized in Proposition 5 of Bowen et al. (2014).

Thus, our experiment is designed to test the status quo effect of mandatory budget rules that institutionalize a relationship between current decisions and future pay-offs, which theoretically leads to efficiency gains. In particular, we propose to test hypotheses that are informed by the equilibrium theory. Since in equilibrium, all proposals are accepted, our data analysis will mainly focus on accepted proposals, though we will also examine factors affecting the acceptance of proposals.⁷

⁶ The two mandatory treatments, aligned or polarized, are distinguished by whether $Y_{DO}^*(high) < \theta_H + \theta_L$ or not, which are named as *low-* or *high-polarization* cases, respectively, in Bowen et al. (2014).

⁷ Bowen et al. (2014), Sec.III, show that any equilibrium is payoff equivalent to the one where (i) responders accept when they are indifferent between accepting and rejecting, and (ii) the equilibrium proposals are always accepted.

Given our design and research questions, we have the following testable hypotheses:

Hypothesis 1 For fixed θ_H and θ_L , public good provision is higher under mandatory budget rules than under discretionary budget rules.

Specifically, under our parameterization, starting from the status quo level of $Y = 1$, public good provision is predicted to grow close to or to achieve the Pareto efficient public good levels under the mandatory budget rules but will remain below this level under the discretionary budget rules.

Hypothesis 2 An increase in the efficient public good provision amount results in an increased steady state allocation to the public good under the mandatory rules.

Under our parameterization, an increase in political polarization (θ_H increasing from 40 to 55) results in a higher level of efficient public good provision. It follows that, as we move from Ma ($\theta_H = 40$) to Mp ($\theta_H = 55$) we should observe greater allocations to the public good by both proposer types eventually. Specifically, even the proposer type whose importance parameter for public good utility doesn't change (the low type in our experiment) has an incentive to increase their public good allocations according to a change in the importance parameter of their opponent type (if the change in the latter parameter results in a change in the Pareto efficient public good amount).

Hypothesis 3 In all treatments, proposers exercise proposer power by generally proposing 0 private points to responders and keeping all points in excess of the public good allocation for their own private accounts.

The testable private point prediction of the model is summarized in the above Hypothesis 3. The theory predicts that in all settings, proposers exercise full proposer power, which means that when the status quo public good allocation is below the Pareto efficient level, any points not allocated to the public good are primarily, if not exclusively, allocated to the proposer's own private account and not to the responder's private account. Figure B.1 in the Appendix shows predictions for private point allocations as a function of the status quo public good level in our parameterization of the two mandatory treatments. When this status quo amount is below the Pareto efficient level, private points allocated to the responder are always zero in the Ma treatment, and sometimes marginally different from zero for certain status quo levels in the Mp treatment (see Figure B.1 in the Appendix for details). Similarly, in the discretionary treatment there is never any allocation of points to the responder's private account. While other bargaining games (e.g., ultimatum bargaining or legislative bargaining) also predict the full exercise of proposer power, a difference here is that there is also a public good component to players' payoffs that benefits both players and thus it is of interest to understand whether or not the presence of this public good component works to strengthen the use of proposer power.

Hypothesis 4 Under discretionary budget rules, in equilibrium both high and low types propose distinct levels for the public good (which are their static equilibrium amounts, $\theta_H = 40$ and $\theta_L = 25$) no matter how many rounds are played in a super-game of the public and private good bargaining task.

As mentioned before, the logic here follows directly from the first order condition for the static optimization. Since an agreement about public good allocations in the current round has no implications for future rounds, the incentives for proposing the static equilibrium levels θ_i for the group account (public good) are still maintained in the dynamic games of the discretionary treatment.

The theory also has predictions regarding the dynamics of behavior and the convergence of public good allocations to steady states over time for the mandatory treatments, which are summarized in the following hypothesis:

Hypothesis 5 Under mandatory rules, starting from the initial status quo level of $Y = 1$ (out of an endowment of 100), both types will propose public good amounts that should converge over time to the steady state levels - that is the high type's dynamic optimum ($Y_{SS}^* \approx 64.615$) in the Ma treatment and the Pareto efficient level ($Y_{SS}^* = 80$) in the Mp treatment.

Finally, we consider some efficiency measures that can be used to evaluate the performance of the different budget rules. As a measure of aggregate efficiency, we look at the ratio of actual payoffs earned from accepted allocations to payoffs that would have been obtained at Pareto optimal allocations. Given the predicted public good allocations across treatments, we have the following:

Hypothesis 6 Aggregate efficiency will be higher in the two mandatory treatments as compared with the discretionary treatment.

The difference in efficiency between the two mandatory treatments is more ambiguous, and we will address this topic later in Sect. 4.9 when we evaluate hypothesis 6.

In the next section we evaluate each of these six hypotheses using the data from our experiment.

4 Experimental results

We report on results from 5 sessions of each of our three treatments, 15 sessions in total. As noted, there are 10 subjects per session; thus we report on data from 150 subjects.

Recall that the discretionary (D) and mandatory-aligned (Ma) treatments only involve a change in the status quo rule for the public good; the values of $(\theta_H, \theta_L) = (40, 25)$, are kept constant between these two treatments. By contrast, the

Table 2 Overview of main outcome variables, all rounds

	Discretionary	Mandatory-aligned	Man- datory- polarized
	(D)	(Ma)	(Mp)
Acceptance rates	88.64%	77.09%	79.10%
Public good allocations	50.81	62.58	71.88
by high type prop	[40]	[64.76]	[84.35]
Private goods to high	34.66	19.47	11.72
type proposers	[60]	[35.17]	[15.07]
Private goods to low	14.53	17.96	16.40
type responders	[0]	[0.07]	[0.58]
Public good allocations	43.33	51.72	57.43
by low type prop	[25]	[45.26]	[51.65]
Private goods to low	48.53	33.81	30.93
type proposers	[75]	[50.79]	[47.82]
Private goods to high	8.15	14.46	11.63
type responders	[0]	[3.95]	[0.53]
Public good allocations	47.16	57.12	64.55
by both type prop	[32.69]	[54.94]	[67.77]
Private goods allocated	41.42	26.69	21.46
to both type prop	[67.31]	[43.03]	[31.67]
Private goods allocated	11.42	16.20	13.99
to both type resp	[0]	[2.02]	[0.55]
Aggregate efficiency	96.02%	97.77%	97.87%

(i) Public good allocations and private point divisions are all as observed in accepted proposals. (ii) Predictions (in square brackets) are based on realized types of proposers/responders and realized status quo levels. (iii) Aggregate efficiency is measured as the ratio of the sum of proposer's and responder's actual payoffs to the same sum of payoffs that would have been achieved at Pareto optimum

mandatory-polarized (Mp) treatment involves both the mandatory rule for the status quo level of the public good and a greater difference between θ_H and θ_L (i.e., greater polarization) namely $(\theta_H, \theta_L) = (55, 25)$ and thus a higher level for efficient public good provision.

As noted earlier, we focus here and throughout the paper on *accepted* proposal amounts in keeping with the theory and since acceptance rates are generally high.

4.1 Overview

We begin with an overview of the main outcome variables from our experiment. Table 2 reports for each treatment, mean values for each of five main outcome variables: (1) proposal acceptance rates, (2) accepted amounts allocated to the public good, disaggregated by high or low type proposers; (3) accepted amounts allocated to the proposers' own private account, disaggregated by high or low type proposers;

Table 3 Mann–Whitney tests of differences in acceptance rates across treatments

High proposer type		Low proposer type		Both proposer types	
Alt.H	<i>p</i> -values	Alt.H	<i>p</i> -values	Alt.H	<i>p</i> -values
D≠Ma	0.009	D≠Ma	0.012	D≠Ma	0.009
D≠Mp	0.076	D≠Mp	0.016	D≠Mp	0.012
Ma≠Mp	0.094	Ma≠Mp	0.917	Ma≠Mp	0.530

p-values for tests of differences in acceptance rates between treatments are reported. The column ‘Alt.H.’ states the alternative hypotheses that acceptance rates between any two treatments are *not* the same (2-sided test). D=Discretionary, Ma=Mandatory-aligned, Mp=Mandatory-polarized

(4) accepted amounts allocated to the responders’ private account, disaggregated by high or low type responders, and (5) aggregate efficiency levels achieved. Means are reported for all rounds of all supergames. The table also shows in square brackets the (Markov perfect) equilibrium predictions based on actual realizations for the proposer types and given the actual status quo levels for the public good that were realized in the experimental games, which is most relevant for the mandatory treatments.

Indeed, as Table 2 reveals, proposals are accepted on average more than 75% of the time. A general finding observed across all treatments is that amounts allocated to the public good by low proposer types are greater than equilibrium predictions, while amounts allocated by high proposer types are lower than equilibrium predictions in the mandatory treatments. On the other hand, both types of *proposers* allocate less, on average, to their own private accounts and more, on average, to the private accounts of their matched responders. Despite these differences, efficiency, as measured by the ratio of total payoffs earned to the Pareto optimum payoff level, is generally quite high, in excess of 95%.

4.2 Acceptance rates

We begin by discussing responder’s acceptance rates of proposals made by proposers. As Table 2 reveals, acceptance rates differ from the equilibrium prediction of 100%, and are highest for the discretionary treatment and lower for the two mandatory treatments. Details on acceptance rates by treatment and session are found in Table B.1 in the Appendix.

The difference in acceptance rates between the discretionary and mandatory treatments likely reflects the fact that under the discretionary rule, the rejection of a proposal means that earnings are zero while under the mandatory rule, if the status quo level of the public good, $Y > 1$, rejection still results in a positive payoff to both players and this status quo payoff level can grow large over time, i.e., the status quo is endogenous. Thus, the mandatory rule gives responders greater bargaining power that gets stronger as the status quo points become higher, and empirically, this results in higher rates of rejection (lower acceptance rates) under the two mandatory rules as compared with the discretionary rule. Indeed, the difference in acceptance

rates between the discretionary treatment and either of the two mandatory treatments (Ma or Mp) is significant at the 5% level as revealed in Table 3 which reports on Mann-Whitney tests using session level mean data (5 sessions per treatment) over various sub-intervals. We observe in Table 2 that the difference in acceptance rates between the discretionary and mandatory treatments is around 10 percentage points, on average. Table 3 reveals that there is *no* difference in acceptance rates between the two mandatory treatments, Ma and Mp. Summarizing this discussion we have:

Result 1 Acceptance rates across all treatments are less than 100%. Acceptance rates are significantly higher in the discretionary treatment as compared with the two mandatory treatments. There are no significant differences in acceptance rates between the two mandatory treatments.

4.3 Determinants of responder acceptance decisions

We next examine the determinants of responder acceptance decisions using a Probit regression analysis. Here the binary dependent variable is equal to 1 if the proposal was accepted and 0 otherwise. The results from our analysis are reported in Table 4 for all treatments and in Table 5 for the two mandatory treatments only.

In Table 4 we observe that responders are more likely to accept offers the higher is the proposer's allocations to the public good and the higher is the proposer's allocation to the responder's private point balance. We further observe that responder acceptance decisions are *decreasing* with increases in the status quo amount of the public good in the two mandatory treatments, Ma and Mp. The latter result follows from the fact that in the discretionary treatment the status quo level is not changing but it generally rises over time in the mandatory treatments. Intuitively, as the status quo level rises, it is easier for responders to reject proposals as the positive status quo level of the public good guarantees that they will get some positive payoffs from the public good (upon rejection). Finally, we observe that controlling for public and private good allocations, the status quo public good level and treatment effects, proposals made by *high* type players (those who value the public good more) are significantly more likely to be rejected by their opponent low type responders. The latter finding is our first indication that *fairness concerns* may play a role in responders' acceptance decisions. We explore such concerns in more detail later in Sect. 4.6 as well as in Appendix C.

These results remain largely robust if we restrict attention to the two mandatory treatments only, as reported on in Table 5. In the analysis of Table 5 we further explore if acceptance decisions depend on whether the status quo level of the public good is below the efficient level (low SQ) or above the efficient level (high SQ). We find that, in a way to facilitate the theory predictions, when the status quo level is below the efficient level, a higher allocation to the public good leads to a greater likelihood of acceptance by responders, and this effect is highly significant. When the status quo level is above the efficient level, higher private points awarded to the responder lead to a small but significant reduction in acceptance rates by responders.

Table 4 Responders' acceptance decisions, all treatments

Variables	y = 1 if a proposal is accepted			
	(1)	(2)	(3)	(4)
Constant	0.197 (0.256)	2.013*** (0.326)	0.140 (0.255)	1.943*** (0.325)
Public good alloc.s	0.025*** (0.002)	0.007** (0.003)	0.025*** (0.002)	0.007** (0.003)
Private points (responder)	0.036*** (0.004)		0.036*** (0.004)	
Status quo	-0.009*** (0.002)	-0.009*** (0.002)	-0.010*** (0.002)	-0.010*** (0.002)
Diff. Private pts (pro.-res.)		-0.018*** (0.002)		-0.018*** (0.002)
Ma	-0.466* (0.281)	-0.466* (0.281)	-0.439 (0.285)	-0.439 (0.285)
Mp	-0.510* (0.290)	-0.510* (0.290)	-0.480 (0.301)	-0.480 (0.301)
High type (proposer)	-0.600*** (0.082)	-0.600*** (0.082)	-0.597*** (0.081)	-0.597*** (0.081)
Sequence	0.005 (0.018)	0.005 (0.018)	0.005 (0.018)	0.005 (0.018)
Round (in sequence)			0.016 (0.017)	0.016 (0.017)
Observations	2985	2985	2985	2985
Pseudo- R^2	0.133	0.133	0.134	0.134

(i) Random-effect probit models are estimated with clustering at the session level for responder's acceptance decisions about proposed allocations. The estimation is based on all proposals (not necessarily accepted). (ii) $Ma=1$ if treatment= Ma ($\theta_H = 40$); $Mp=1$ if treatment= Mp ($\theta_H = 55$). (iii) Diff. Private pts.=Difference between proposer's and responder's private points. (iv) Standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. (v) $Ma - Mp = 0.044$ (0.1352), 0.044 (0.1352), 0.041 (0.138), 0.041 (0.1385) in columns (1)-(4), respectively, are all *not* statistically significant (standard errors in parentheses). (vi) We use McFadden's Pseudo $R^2 = 1 - \ln(L_M)/\ln(L_0)$, where $\ln(L_M)$ and $\ln(L_0)$ are the log likelihood from the full model and the model only with constant, respectively

We further consider differences in mean allocations to public good and private points between accepted and rejected proposals - see Table B.2 and Figure B.2 in the Appendix. There we show that public good and responder private points are significantly greater in accepted proposals as compared with rejected proposals while proposer private points are significantly lower in accepted proposals as compared with rejected proposals. This evidence further confirms that responders consider both public good levels and their own private points in deciding whether to accept proposals, as was already shown in the probit regression results.

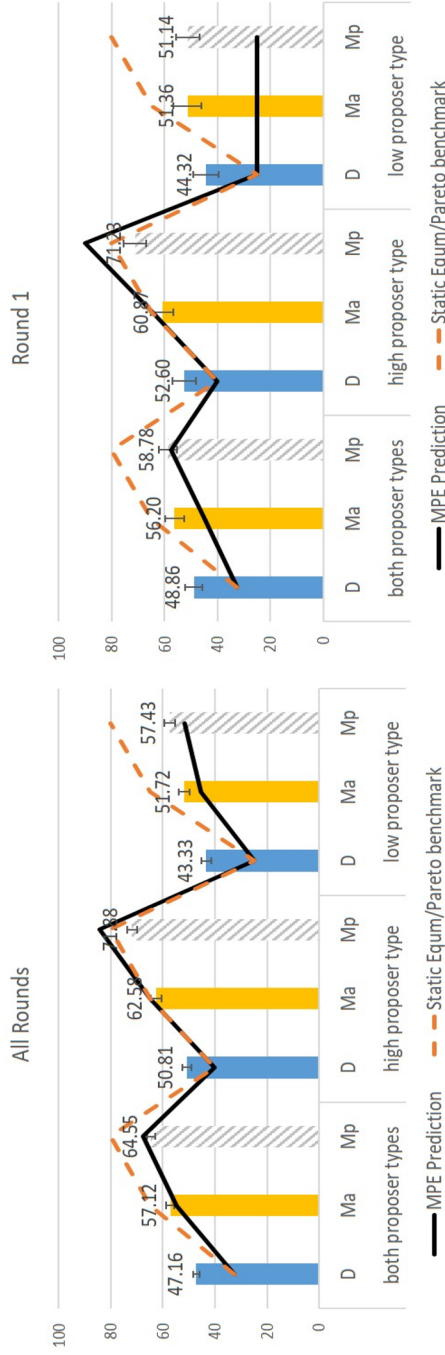
Table 5 Responders' acceptance decisions, mandatory treatments only

Variables	y = 1 if a proposal is accepted		
	(1)	(2)	(3)
Constant	-0.394** (0.191)	1.561*** (0.220)	-0.707*** (0.128)
Public good alloc.s	0.028*** (0.001)	0.009*** (0.003)	0.022*** (0.002)
Private points (responder)	0.039*** (0.005)		0.040*** (0.006)
Status quo, SQ	-0.010*** (0.002)	-0.010*** (0.002)	
Diff. Private pts (pro.-res.)		-0.020*** (0.002)	
Mp	-0.070 (0.142)	-0.070 (0.142)	-0.158 (0.128)
High type (propoeer)	-0.613*** (0.086)	-0.613*** (0.086)	-0.577*** (0.101)
Sequence	-0.004 (0.018)	-0.004 (0.018)	
Round (in sequence)	-0.009 (0.015)	-0.009 (0.015)	
Public×low SQ			0.006*** (0.001)
Private×high SQ			-0.014*** (0.004)
Observations	1990	1990	1990
Pseudo- R^2	0.150	0.150	0.137

(i) Random-effect probit models are estimated with clustering at the session level for responders' acceptance decisions about proposed allocations. The estimation is based on all proposals (not necessarily accepted). (ii) Mp=1 if treatment=Mp ($\theta_H = 55$). (iii) Diff. Private pts.=Difference between proposer's and responder's private points. (iv) Standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. (v) Low SQ (dummy) indicates the cases when the status quo public good is below Pareto efficient level while high SQ (dummy) indicates the opposite case. (vi) We use McFadden's Pseudo $R^2 = 1 - \ln(L_M)/\ln(L_0)$, where $\ln(L_M)$ and $\ln(L_0)$ are the log likelihood from the full model and the model only with constant, respectively

4.4 Effect of mandatory rules on public good levels

We now consider the main treatment effect of adopting mandatory budget rules for public good provision relative to the discretionary budget rule case. We first focus on a comparison of the mandatory-aligned (Ma) and discretionary (D)



Notes. D=Discretionary, Ma=Mandatory-aligned, Mp=Mandatory-polarized. Error bars indicate 95% confidence intervals.

Fig. 3 Mean Accepted Allocation to Public Good, All Sequences. Left panel: all rounds; Right panel: the first round.

Table 6 Nonparametric tests for differences in (accepted) public good allocations

	All Rounds		Round 1	
	Alt.H	<i>p</i> -values	Alt.H	<i>p</i> -values
High	D<Ma	0.008	D<Ma	0.059
P.Type	Ma<Mp	0.014	Ma<Mp	0.008
Low	D<Ma	0.038	D≠Ma	0.251
P.Type	Ma<Mp	0.087	Ma≠Mp	0.917
Both	D<Ma	0.024	D<Ma	0.059
P.Types	Ma<Mp	0.008	Ma<Mp	0.301
D	Low<High	0.059	Low<High	0.022
Ma	Low<High	0.014	Low<High	0.022
Mp	Low<High	0.005	Low<High	0.022

p-values for tests of differences in accepted public good offers between two different treatments or between two different types of proposers within a treatment are reported. The columns 'Alt.H.' state the alternative hypotheses that the public good offers in the 1st treatment are less than those in the 2nd treatment or that the public good offers by low type proposers are less than those by high type proposers in each of the 3 treatments (1-sided test); the predicted public good amounts for the low type in round 1 are the same across all 3 treatments (2-sided test). D=Discretionary, Ma=Mandatory-aligned, Mp=Mandatory-polarized. P.Type=proposer type

treatments as they are most comparable (have the same θ values). We report the following finding.

Result 2 Consistent with Hypothesis 1, public good provision is higher in the mandatory-aligned treatment than in the discretionary treatment.

Support for Result 2 comes from Fig. 3 which reports mean public good allocations by treatment and proposer type over all rounds and for the first rounds of sequences. Further Table 6 provides results from non-parametric tests on accepted public good amounts by treatment over all rounds or round 1 only using session level averages (session level data on public good allocations reported on in Tables B.3-B.5 of the Appendix). In Table 6, and those that follow (about non-parametric tests), we used one-sided tests whenever we have specific directional predictions from the theory and two-sided tests otherwise.

We first compare allocations to the public good in the discretionary treatment (D) with the mandatory-aligned treatment (Ma). Since the θ values do not change between these two treatments, the comparison of D vs. Ma provides the cleanest test of the effect of changing the bargaining rules. We observe that over all rounds of all sequences, the mean agreed upon public good allocation proposed by both high and low types in the Ma treatment is around 10 points higher than the mean

Table 7 Tobit regression analysis of accepted public good allocations, all treatments

Variables	Accepted public good allocations		
	(1)	(2)	(3)
Constant	40.462*** (1.969)	47.468*** (1.948)	47.735*** (2.140)
Private Points (responder)	-0.641*** (0.029)	-0.654*** (0.029)	-0.634*** (0.029)
Ma	13.143*** (2.399)	9.494*** (2.702)	4.739 (3.000)
Mp	20.330*** (2.395)	15.621*** (2.696)	11.946*** (2.981)
High Type (proposer)	15.304*** (0.756)	15.117*** (0.760)	15.221*** (0.755)
Sequence	0.904*** (0.160)		-0.195 (0.264)
MaxSeq			1.770*** (0.386)
Mp×Seq			1.775*** (0.384)
Round (in sequence)	0.564*** (0.137)	-0.094 (0.223)	
MaxRd		0.980*** (0.331)	
Mp×Rd		1.247*** (0.332)	
Observations	2436	2436	2436
Pseudo- R^2	0.036	0.035	0.037

(i) Random-effect tobit model is estimated with lower bound=1 and upper bound=100 in the dependent variable. The estimation is based on accepted proposals. (ii) Ma=1 if treatment=Ma ($\theta_H = 40$); Mp=1 if treatment=Mp ($\theta_H = 55$). (iii) Standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. (iv) We use McFadden's Pseudo $R^2 = 1 - Ln(L_M)/Ln(L_0)$, where $Ln(L_M)$ and $Ln(L_0)$ are the log likelihood from the full model and the model only with constant, respectively

for the discretionary treatment and this difference is significant at the 5% level using the Mann–Whitney test on session level data as revealed in Table 6.

Using a random-effects Tobit regression analysis (to account for data censoring) and *all* data on accepted public good amounts from all treatments, Table 7 confirms that public good provision is, on average, significantly higher in the

Table 8 Tobit regression analysis of accepted public good allocations, mandatory treatments only

Variables	Accepted public good allocations		
	(1)	(2)	(3)
Constant	49.429*** (1.824)	53.350*** (1.727)	49.077*** (1.984)
Private points (responder)	-0.750*** (0.036)	-0.766*** (0.036)	-0.752*** (0.036)
Status quo	0.199*** (0.016)	0.216*** (0.016)	0.190*** (0.014)
Mp	6.158*** (1.812)	6.346*** (2.172)	6.055** (2.524)
High type (proposer)	15.115*** (0.882)	15.071*** (0.890)	15.117*** (0.883)
Sequence	0.921*** (0.194)		0.920*** (0.273)
Mp×Seq			0.031 (0.378)
Round (in sequence)	-0.217 (0.190)	-0.227 (0.248)	
Mp×Rd		-0.100 (0.331)	
Observations	1554	1554	1554
Pseudo- R^2	0.062	0.060	0.062

(i) Random-effects Tobit model is estimated with a lower bound=1 and upper bound=100 in the dependent variable. The estimation is based on accepted proposals. (ii) Mp=1 if treatment=Mp ($\theta_H = 55$). (iii) Standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. (iv) We use McFadden's Pseudo $R^2 = 1 - Ln(L_M)/Ln(L_0)$, where $Ln(L_M)$ and $Ln(L_0)$ are the log likelihood from the full model and the model only with constant, respectively

Ma treatment as compared with the baseline D treatment in most specifications including those with round and sequence numbers.

We note further that in the discretionary treatment, Table 2 reveals that there is considerable *over*-allocation to the public good in that mean accepted public good amounts are greater than theoretical predictions.⁸ We see considerably less over-allocation in the mandatory-aligned (Ma) treatment.

We further observe in Fig. 3 and in Table 7 that, consistent with the theory, average allocations to the public good in the mandatory-polarized treatment (Mp) are also significantly greater than in the discretionary treatment (D) by somewhere

⁸ Note that the predicted public good amounts (in square brackets) for the mandatory treatments in Table 2 and other tables *condition on the realized status quo level of the public good* at time a proposal is made, which differs by round and across sessions.

between 12–20 points, though in this case, the additional change in the degree of polarization between the D and Mp treatment is a confounding factor.

We next consider the impact on allocations to the public good under the two mandatory rules when there is an increase in polarization, that is, we make a comparison between mean allocations by low and by high types in the mandatory-aligned (Ma) and the mandatory-polarized (Mp) treatment. We have:

Result 3 Consistent with Hypothesis 2, when there is an increase in polarization (hence a change in treatments from Ma to Mp in the experiment), then under the mandatory rules, both types increase their allocations to the public good.

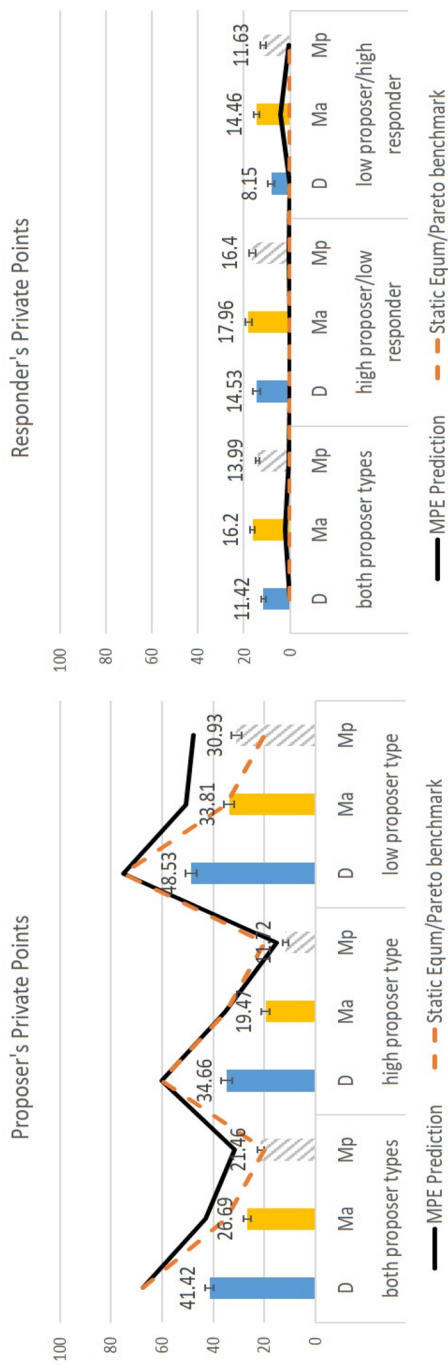
Support for Result 3 comes from Fig. 3 and Tables 6 and 8. From Fig. 3 we observe that average allocations to the public good are higher for both high and low types in the mandatory-polarized (Mp) treatment relative to the allocations of these same types in the mandatory-aligned (Ma) treatment. Using Mann–Whitney tests on session level averages over all rounds, as reported on in Table 6 we find that this difference is significant for either types at the 5% or 10% level of significance. That is, for each type (high or low), we reject the null hypothesis that mean allocations to the public good are the same in treatments Ma and Mp in favor of the alternative that mean allocations are higher in Mp as compared with Ma. Finally, Table 8, reports on another Tobit regression analysis for accepted public good allocations but for the mandatory treatments only and confirms this finding. Across several different specifications, we see that accepted public good allocations are significantly higher in the Mp treatment as compared with the baseline Ma treatment at the 1% or 5% significance level.⁹

Finally, we also examined mean first round choices over all sequences, since in the first round, the status quo level is the same across the three treatments - see the right panel of Fig. 3. There we see that high types across the two mandatory treatments proposed higher allocations to the public good in round 1 than did low types which suggests that high types understood and acted upon the insurance role. Further, high types were especially responsive to changes in the bargaining rule, by monotonically increasing their round 1 allocation to the public good as the bargaining rule changed from D to Ma to Mp. Finally, as Table 6 shows, these differences in round 1 behavior are often significant.

4.5 Accepted allocations to private accounts of proposer and responder

Thus far, we have focused on accepted allocations to the *public good* benefiting both players. However, it is also of interest to consider the amounts allocated to the proposer and responder's *private accounts*. Here again we focus on *accepted* proposals made by the proposer to his/her own private account and to the opponent's

⁹ The same tobit models in Tables 7–8 are estimated separately for each of high and low type proposers and the results are shown to be robust in the subsamples of fixed proposer types in Appendix Tables B.6–B.7.



Notes. D=Discretionary, Ma=Mandatory-aligned, Mp=Mandatory-polarized. Error bars indicate 95% confidence intervals.

Fig. 4 Mean Accepted Allocation to Private Account, All Sequences. Left panel: proposer; Right panel: responder.

Table 9 Nonparametric tests for differences in accepted private points to proposer and responder

Private points (Proposer)			Private points (Responder)		
	Alt.H	<i>p</i> -values		Alt.H	<i>p</i> -values
High	D>Ma	0.014	D	Low>High	0.040
P.Type	Ma>Mp	0.005	Ma	Low>High	0.069
Low	D>Ma	0.059	Mp	Low>High	0.112
P.Type	Ma>Mp	0.174	Pool	Low>High	0.005
Both	D>Ma	0.024			
P.Types	Ma>Mp	0.014			

p-values for tests of differences in accepted proposers' private points between 2 different treatments or in accepted responders' private points between 2 different types within the same treatment. The columns 'Alt.H.' state the alternative hypotheses that the proposers' private points in the 1st treatment are *greater* than those in the 2nd treatment, e.g., D>Ma, or that the private points assigned to low type responders are *greater* than those assigned to high types in each of the 3 treatments or all treatments pooled together, Low>High (1-sided test). D=Discretionary, Ma=Mandatory-aligned, Mp=Mandatory-polarized. P.Type=proposer type

(responder's) private account. Mean amounts for both allocations are reported on in Table 2 and illustrated in Fig. 4.

Figure 4 and the non-parametric tests reported in Table 9¹⁰ reveal that, consistent with theoretical predictions, private points allocated to proposers are significantly higher in the discretionary treatment as compared with the Ma treatment, where they are significantly higher than in the Mp treatment. We further note that accepted private points to low type responders are significantly greater than accepted private points to high type responders.

However, in all cases accepted points allocated to the proposer's private accounts lie below the predicted amounts based on the realized status quo level (solid line) or using the Pareto efficient equilibrium benchmark (dashed line) in Fig. 4. That is, proposers of both types *under*-allocate to their own private accounts on average and they *over*-allocate to the responder's private account on average, relative to theoretical predictions.

While the equilibrium suggests that a proposal will be accepted so long as there is sufficient provision of public goods but zero private points to responders, especially given a status quo below the Pareto efficient level, such proposals are typically rejected in our laboratory experiment and proposers had to offer private points as well to their responders at most realized status quo levels. Indeed, as Tables B.15-B.16 in the Appendix show, mean proposed amounts to the proposer's private account and to the responder's private account, independent of the acceptance decision are, respectively, slightly greater and lower than are the same amounts

¹⁰ Based on the session level data in Appendix Tables B.8-B.10 for proposers and Appendix Tables B.11-B.13 for responders.

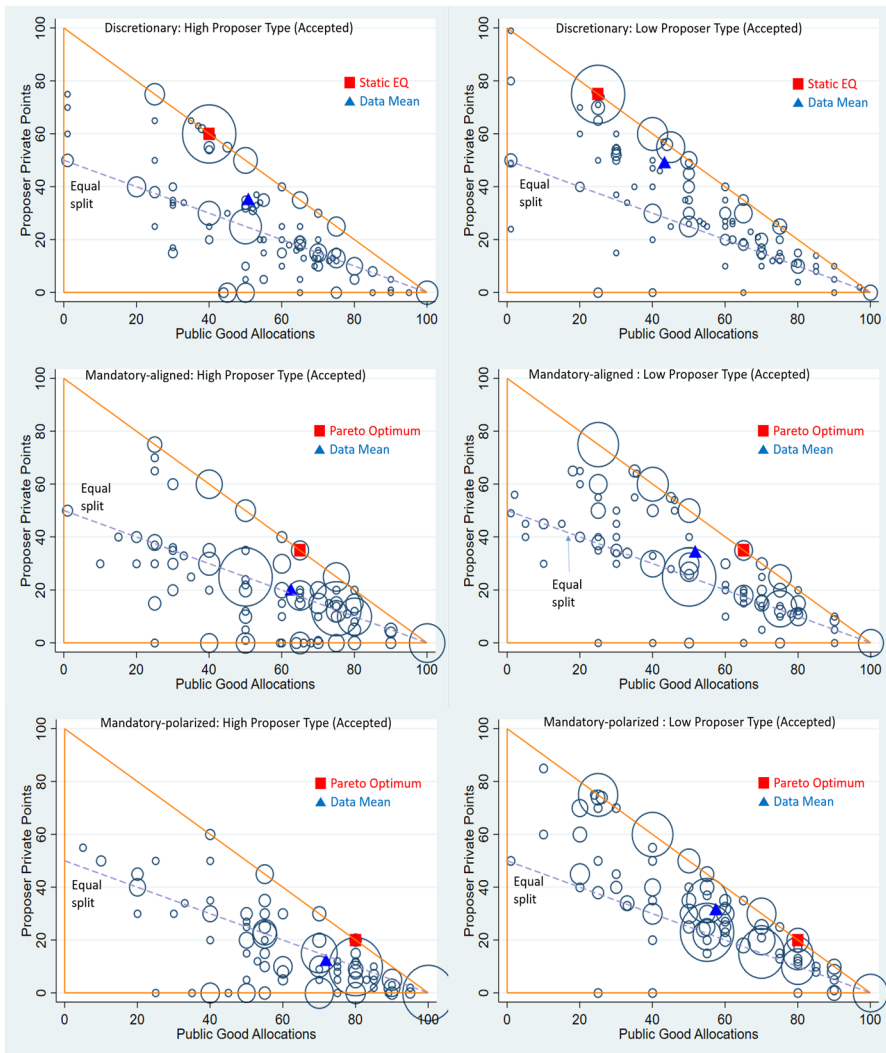


Fig. 5 Bubble plots of allocation vectors by proposer types and treatments, accepted proposals only

conditional on acceptance of the proposal in Appendix Tables B.8 and B.11.¹¹ We understand this as proposers not being able to fully exercise their *proposer power*, a widely observed phenomenon in the empirical bargaining literature.

¹¹ Proposed public good amounts, independent of acceptance, are also shown in Appendix Table B.14.

4.6 Allocations within the 2D simplex

Figure 5 summarizes, using a 2-dimensional simplex, the frequency of *accepted* allocation vectors made by high and low proposer types. The advantage of this approach relative to our analysis thus far is that, instead of looking at one-dimensional analysis of public or private points, here we can consider the behavior of allocation vectors in multidimensional (bargaining) choice spaces. In each panel of the figure, the first coordinate, on the horizontal axis, is the amount allocated to the public good and the second coordinate, on the vertical axis, is the proposer's allocation to his own private account. The allocation to responders' private account is the residual amount. Thus, the coordinate pair (50,25) corresponds to case where the proposer allocated 50 points to the public good, 25 points to his own private account and the remaining 25 points went to the responder's private account.¹² The size of a bubble centered at each observed coordinate pair is proportional to the count of observations with that allocation vector; the smallest bubble corresponds to a single observation. These figures also show the mean accepted allocation from the experimental data (Data Mean, indicated by the solid triangles) along with the static equilibrium prediction in the discretionary treatment and the Pareto optimum allocation in the two mandatory treatments (the solid squares) for reference purposes.

One key finding from these simplex figures is that most *accepted* proposals do *not* lie on the hypotenuse of the simplex triangle where the proposer exercises full proposer power, keeping all points not allocated to the public good for himself. Instead, the most frequently observed accepted proposals involve an equal division of private points between the proposer and the responder. Such "equal split" allocations are defined as those for which the proposer divides the amount *not* allocated to the public good equally between himself and the responder. These equal split allocations are found along the dashed line labeled "equal split" in Fig. 5. While we observe these equal split outcomes in all three treatments, the frequency of such equal split allocations is greater in the two mandatory treatments.

A second key finding concerns the mean accepted allocations relative to theoretical predictions. We see a large difference in public good allocations between the discretionary treatment and the two mandatory treatments which is consistent with theoretical predictions (and earlier findings). In the discretionary treatment, the upper row of Fig. 5, we observe a large mass of observations around the static equilibrium allocation. By contrast, as we move from the discretionary to the two mandatory treatments, the middle and bottom panels, we see an increase in the accepted amounts allocated to the public good, as indicated by the rightward movement of the data mean allocation. This movement is away from the static equilibrium and toward the Pareto efficient outcome. The shift is particularly pronounced for high proposer types and only less so for low proposer types. In Appendix Figures B.3-B.5, we provide further evidence of similar movements over time, between the first and second

¹² Also note that the three vertices of the 2-dimensional simplex in Fig. 5 represent proposals: that allocate all available resources to the group account (100, 0); that allocate all resources to the proposer's private account (0, 100); or that allocate all resources to the responder's private account (0, 0).

halves of sessions, in the mean accepted allocations in all three treatments. In the discretionary treatment there is a very slight movement toward the static equilibrium while in the two mandatory treatments there is a more pronounced movement toward the Pareto optimum allocation over time.

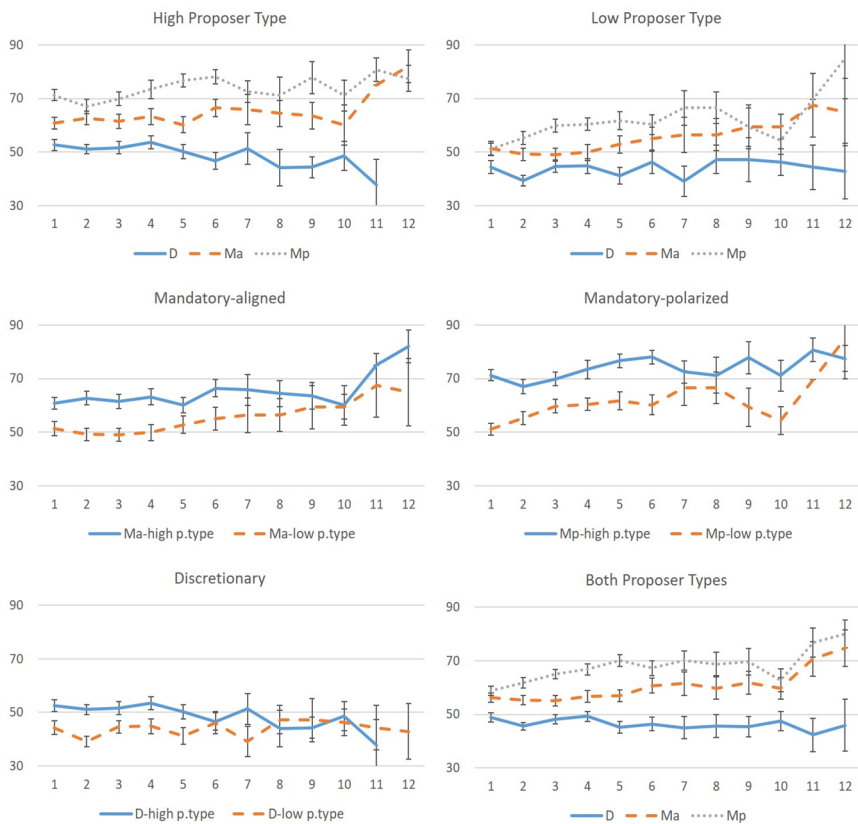
A further observation from the 2D simplex Fig. 5 is that in the two mandatory treatments, low type proposers' accepted allocations assign greater points to their own private accounts than do high type proposer's accepted allocations. Low type proposer's accepted allocations generally lie *on or above* the equal split line, while high type proposer's accepted allocations generally lie *on or below* the equal split line. The difference is largely due to some high types proposing allocations at the *lower* bound of the simplex (the horizontal leg of the triangle) where high type proposers are giving *all* of the endowment points in excess of the public good to the responder. We seldom see this type of allocation behavior in the case of low type proposers suggesting that fairness motivations are playing an important role since low type *responders* (matched with high type proposers) don't get the same benefit from the public good as do high type *responders* (matched with low type proposers).

In the two mandatory treatments, it may be puzzling to find that low type proposers are proposing public good amounts that are substantially below the Pareto optimum level, as is also revealed in Fig. 3. Figure 5 suggests that low types in the two mandatory treatments appear to be allocating too much, on average, to their own private accounts, particularly in the Mp treatment (see also Fig. 4 for low types' over-allocation to their own private accounts, relative to the Pareto benchmark - drawn as the dashed line). However, Figs. 3 and 4 also reveal that low type proposers have actually over-allocated to public goods and under-allocated to their own private account in the two mandatory treatments, taking into account the predictions based on the realized status quo levels of the public good (i.e., the predictions depicted as the solid line). Thus, it seems that a reason for the under-allocation to the public good by the low proposer types is that the status quo level of the public good is not high enough in the time frame of our experiment to justify their allocating the Pareto efficient amount to the public good.¹³ Of course, the status quo level depends on the behavior of both high and low type proposers. The high proposer types in the mandatory treatments are generally closer to the Pareto optimal public good levels, particularly in the Ma treatment.

Based on Fig. 5, we summarize our findings regarding proposer power in relation to Hypothesis 3 as follows:

Result 4 Inconsistent with Hypothesis 3, proposers do not exercise full proposer power by allocating zero private points to responders. This is particularly evident in the mandatory treatments, where sizeable fractions of proposers divide points net of the public good allocation equally between their own private accounts and those of responders.

¹³ Indeed, we find that that status quo public good levels faced by low proposer types are lower than those faced by high proposer types in both mandatory treatments (e.g., this finding can be seen when Appendix Table B.19 is estimated separately for each proposer type in the two mandatory treatments).



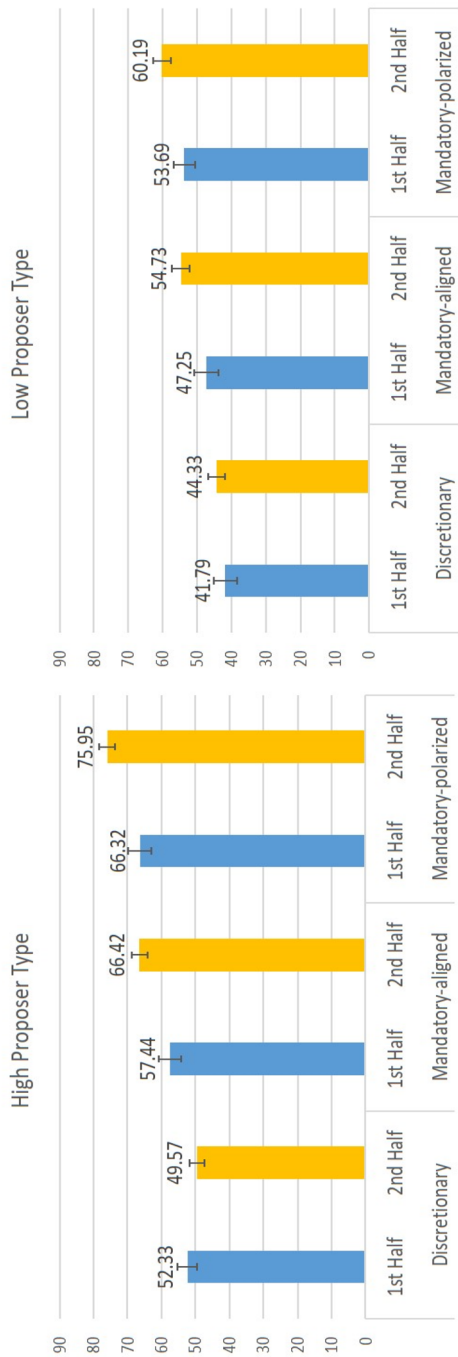
Notes. D=Discretionary, Ma=Mandatory-aligned, Mp=Mandatory-polarized. p.type=proposer type. Error bars show standard errors. Horizontal axis indicates rounds in a sequence (supergame).

Fig. 6 Mean (Accepted) Public Good Allocations over Rounds.

Given the observed heterogeneity in the extent to which proposers exercise their proposer power, we further explore, in Appendix C, the possibility of behavioral model explanations, using either the Fehr and Schmidt (1999) specification of *other-regarding* preference or a finite mixture model (FMM; see Moffatt (2016)). The latter model gives, for instance, the estimated proportion of *equal splitters*.

4.7 Evolution of accepted public good amounts over time

In this section we explore in further detail, the evolution of agreed upon public good allocations over time. Figure 6 shows mean public good allocations over all rounds of a supergame, with standard error bars (see also Table B.17 in the Appendix). Note that due to our use of random termination to implement a discount factor of $\delta = .80$, earlier rounds in a supergame will have more observations than later rounds



Notes. 1st Half=Sequences (supergames) 1-4, 2nd Half=Sequences 5-7,8,9. Error bars indicate 95% confidence intervals.

Fig. 7 Mean Accepted Allocation to Public Good by Proposer Type, 1st vs. 2nd Half of Session. Left panel: high type; Right panel: low type.

(so there are accordingly, larger standard errors in later rounds) and that the longest supergame in any session was 12 rounds. The top panel of Fig. 6 shows mean (accepted) public good allocations by proposer type, high or low, across all three treatments. The two middle panels and the bottom left panel show mean public good allocations by high and low types for each of the three treatments, Ma, Mp and D. Finally, the bottom right panel shows mean public good allocations for both proposer types combined, across all three treatments.

The clear impression given by Fig. 6 is that over the course of a supergame there is on average, good separation in mean public good amounts across treatments, offered by the high and low proposer types. Further, in both of the mandatory treatments we observe an upward trend in public good allocations while in the discretionary treatment we see a constant or even a declining trend in mean public good amounts.

We next consider each treatment in turn, beginning with the discretionary treatment. Recall from Hypothesis 3 that for the discretionary treatment, high and low types should simply propose their static equilibrium amounts in each round that they serve as the proposer, specifically $Y = \theta_H = 40$ for high types and $Y = \theta_L = 25$ for low types, ignoring any dynamic aspects of the repeated game. As Table 2 and Fig. 3 reveal, in the discretionary (D) treatment, accepted proposals made by high types average 50.81 while those made by low types average 43.3 over all rounds. These levels are *greater* than the static equilibrium levels of 40 and 25 respectively. However, as Table 6 reveals, we can reject the null hypothesis that accepted public good offers by low types are the same as accepted public good offers by high types in favor of the alternative that the latter offers by high types are significantly greater than the former offers by low types at the 10% level of significance ($p = .059$). Further, as Fig. 7 reveals, there is not much change in the accepted public good allocations proposed by high and low types over the first and second halves of each session; that is, while accepted amounts proposed by both types are greater than the static equilibrium levels, they are not increasing or decreasing by much over time. A similar observation follows from the bottom left panel of Fig. 6. We summarize this finding as follows.

Result 5 Consistent with Hypothesis 4 in the discretionary treatment, accepted public good proposals by high types are greater than accepted public good proposals by low types, and do not change much over time. Both types' accepted public good proposal amounts are greater than the static equilibrium levels $(\theta_H, \theta_L) = (40, 25)$.

We next go back to Table 7 and compare the evolution of accepted public good amounts over time in the mandatory treatments relative to the discretionary treatment using Tobit regressions that account for data censoring. Specifically, we report on random-effect Tobit regressions and we include round and sequence numbers to consider behavior over time. The dependent variable is accepted public good amounts within the implemented limits between 1 and 100.

In the regressions reported in Table 7, the discretionary treatment serves as the baseline. We see in specification (1) that the baseline accepted public good

amount is about 40 and is increasing in the mandatory-aligned (Ma) treatment by about 13 and increasing further in the mandatory-polarized (Mp) treatment by around 20. Further, the inclusion of sequence and round numbers in specification (1) suggests that allocations to the public good are growing over time. However, disaggregating this effect further using interaction variables, MaxRound and MpxRound in specification (2) and MaxSequence and MpxSequence in specification (3) we see that the growth in public good allocations over time is owing to the two mandatory treatments (Ma and Mp); including the interactive terms, the coefficients on the round or sequence number variables for the baseline discretionary treatment are no longer significantly different from zero. This finding is consistent with the theory which predicts that accepted public good amounts should be growing over time in the mandatory treatments due to the role played by the status quo public good level in dynamic bargaining under mandatory rules. These same results generally continue to hold if we disaggregate accepted public good allocations by the *type* of player (high or low) who made the proposal as reported on in Appendix Table B.6 (with stronger effects of learning by high types).

Figure 7 also clearly shows that proposers of both types learn to increase public good allocations significantly in the two mandatory treatments as they move from the first to the second half of sessions.

4.8 Adjustment in public good allocations conditional on status quo level

We next look for more explicit evidence of dynamic adjustment *within* a sequence (indefinitely repeated game) in the two mandatory treatments, since in those treatments, each new sequence starts with a status quo level for the public good reset to the initial condition, $Y = 1$, and then, depending on whether proposals are accepted or not, the status quo level for the public good can increase over time.

Looking at the data from the first and second half of sessions (see Table B.1), we observe that acceptance rates remain roughly constant over time and consistently below the 100% equilibrium prediction across treatments (the 95% confidence intervals of mean acceptance rates - viewed as sample proportions - were *overlapping* between the first and the second half of sessions for all three treatments).

We notice further in Fig. 7 that both high and low types tend to increase their public good allocations over time in the two mandatory treatments, while as previously noted, there is not much change in public good allocations over time in the discretionary treatment (the 95% confidence intervals of mean public good allocations in accepted proposals were *non-overlapping* between the first and the second half of sessions for the two mandatory, Ma and Mp, treatments, for each of the two types, high and low, of proposers, while the same intervals were *indeed* overlapping between the two halves for either type of proposers in the discretionary treatment). The latter observation is consistent with the notion that players are learning to play according to the dynamic equilibrium predictions of the theory with greater experience.

Figure 8 shows scatter plots of accepted public good amounts as a function of the status quo level at the time the proposal was made. In addition, we show the fit of

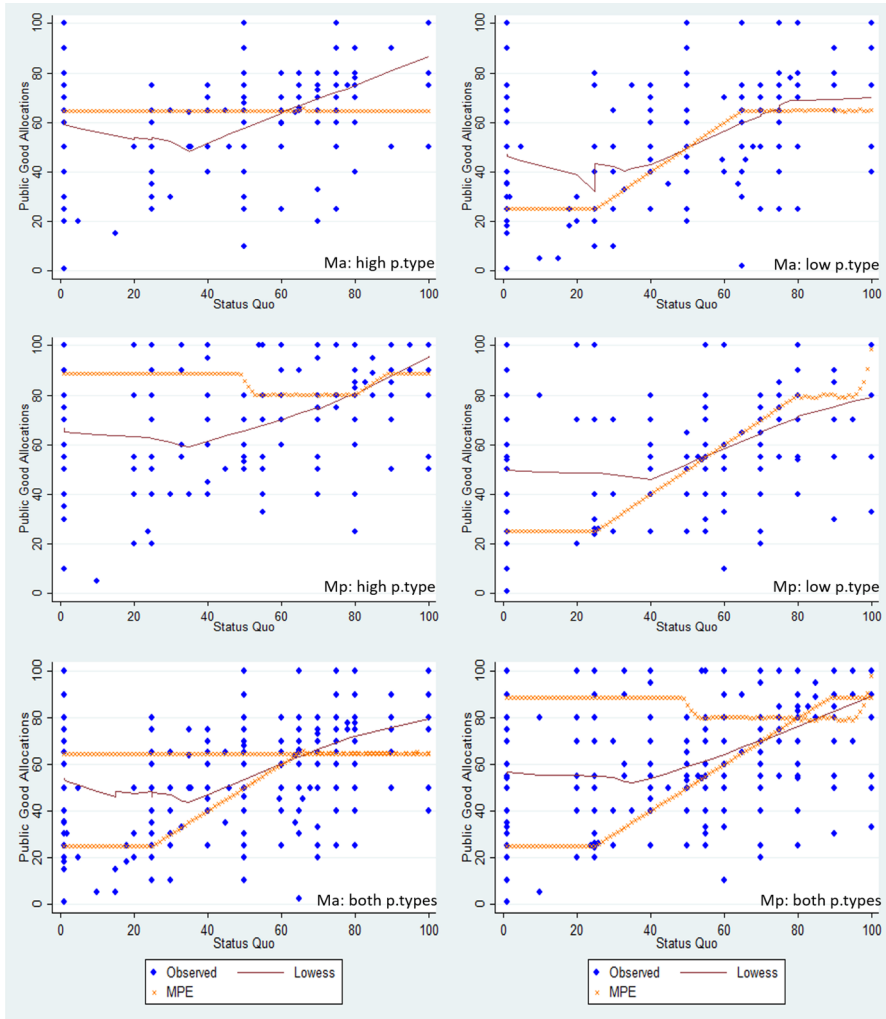


Fig. 8 Scatter plot of (accepted) public good allocations by status quo default level, mandatory treatments along with lowess filter and MPE prediction

Lowess filters to these data. These fitted lines can be compared with the Markov Perfect Equilibrium (MPE) predictions shown together (which are repetitions of those in Fig. 2). The top panels are for the high and low proposer types in the mandatory-aligned (Ma) treatment, the middle panels are for the high and low types in the mandatory-polarized (Mp) treatment, and the bottom panels consider both proposer types combined in Ma (left) and Mp (right). By comparison with the theoretical predictions we find both differences and similarities. On the one hand, we observe that for the low types (the right columns of the top and middle panels), the pattern of public good allocations as a function of the status quo level is qualitatively similar to

the MPE path. On the other hand, for the high types (the left columns of the top and middle panels) public good allocation levels should be more or less constant according to the equilibrium while the data shows a clearly increasing pattern as status quo levels increase.

By way of an explanation, we note that the lowest status quo value ($Y = 1$) is more likely to be observed, as it is the initial state of all supergames, and there is a wide variance of public good allocation amounts for this status quo level for both high and low proposer types. This large initial variance reflects some initial learning/coordination that the theory does not address. Further, as the probit regressions in Tables 4 and 5 revealed, proposals by high type proposers are significantly more likely to be rejected by low type responders across all treatments (while the MPE, as depicted in Fig. 2 and Appendix Figure B.1, assumes no rejection). The greater rejection of high type proposals may cause these high type players to increase their allocations to the public good in order to gain acceptance.

To better understand the repeated game dynamics *within* a supergame (or sequence), we consider a simple first order autoregressive model of the convergence behavior of different outcome variables, which we label y . Specifically, we consider the model

$$y_{j,t} = \lambda y_{j,t-1} + \mu_j + \epsilon_{j,t} \quad (1)$$

where $y_{j,t}$ denotes the time t value of variable j . We are particularly interested in two main outcome variables, namely the accepted amount of the public good in period t , PG_t , and the status quo level for the public good in period t , SQ_t , in the two mandatory treatments, as both of these variables are expected to converge to steady states over time.

Provided that estimates of λ are less than 1, the steady state public good and status quo amounts over a supergame are well approximated by estimates of the limiting value of equation (1), namely, by $\frac{\mu}{1-\lambda}$. Estimation of equation (1) for the accepted public good amount (PG) are reported on in Appendix Table B.18 while estimates for the status quo level of the public good (SQ) are reported on in Appendix Table B.19.

These tables reveal several things. First estimates for λ are generally less than 1 providing evidence of weak convergence over time across all treatments. Second, the limiting estimated values for $\frac{\mu}{1-\lambda}$ in the two mandatory treatments are greater than in the discretionary treatment and are close to, but often fall just short of predicted steady state values for these two mandatory treatments. For instance, considering both proposer types in treatment Ma (Ma-both) a 95% confidence interval for the estimated limiting value of the public good allocation or the status quo level does not include the steady state level of 64.615, though the data are very close to this level. For high type proposers in treatment Ma (Ma-high), the 95% confidence interval for the estimated limiting value of the public good allocation overshoots the predicted steady state level of 64.615. Similarly, for the Mp treatment, considering both proposer types (Mp-both), a 95% confidence interval for the estimated limiting value of the public good allocation or the status quo level does not include the steady state level of 80, though again the data are very close to this level. For high

proposer types in Mp (Mp-high), the 95% confidence interval for the estimated limiting value of the public good allocation does include the predicted steady state level of 80 (for high types in Ma, the same intervals can even stay above the predicted level of 64.615). A third finding from Appendix Tables B.18-B.19 is that the 95% confidence intervals for the estimated limiting values of accepted public good allocations and status quo levels are non-overlapping across all three treatments (when both proposer types are combined, and in the second half and over all rounds of session); that is, there is *good separation* in these limiting values as we move from the discretionary treatment to treatment Ma and then to treatment Mp. For instance, considering both proposer types and all rounds, Appendix Table B.18 reveals that the estimated limiting accepted public good allocation in the discretionary treatment (D-both) is 46.688 with a 95% confidence interval of [44.455, 48.921]; for the Ma treatment the estimated limiting accepted public good allocation for both proposer types (Ma-both) is 59.749 with a 95% confidence interval of [56.780, 62.718]; and finally for the Mp treatment the estimated limiting accepted public good allocation for both proposer types (Mp-both) is 69.537 with a 95% confidence interval of [66.756, 72.318]. We summarize these findings as follows:

Result 6 Regarding Hypothesis 5, in the mandatory treatments, accepted public good allocations are converging toward steady state levels, but estimated limits often fall just short of steady state predictions. Still, there is good separation of the long-run mean public good allocations across treatments in terms of 95% confidence intervals. A similar pattern obtains for the convergence of the status quo public good level across treatments.

4.9 Efficiency

In this section we consider efficiency across all three treatments. To enable a consistent comparison, we calculate efficiency as the ratio of actual payoffs from accepted allocations to those that would have been obtained at the Pareto optimal allocations as was already done in the overview Table 2. While the Pareto optimum allocation is not an equilibrium under the discretionary rules, using the Pareto optimum payoff levels as a benchmark enables comparisons across all three treatments. Here we report both *aggregate* efficiency measures, defined as the sum of proposers' and responders' actual payoffs relative to the Pareto optimum payoff level, and *individual* type-specific payoffs: proposer/responder and high/low types payoffs relative to the Pareto optimum. To be precise, the denominator of both the aggregate and individual efficiency measures uses the same (hypothetical) aggregate payoff at the Pareto optimum since we are interested in how aggregate efficiency is decomposed individually between high and low type proposers and responders. The results are shown in Table 10 and nonparametric tests for differences in these efficiency measures are reported in Table 11.¹⁴

¹⁴ Table B.20 in the Appendix presents an alternative version of individual efficiency where the denominator is the individual type/role's portion of the Pareto optimal payoff. In matches between high proposers and low responders, the low types get more than their own share of Pareto optimum payoffs while the high types get less, which reflects fairness concerns but contrasts with the original efficiency measures

Table 10 Aggregate and Individual Efficiency

	Efficiency (%)	PO prediction (%)
D: Aggregate-both p.types	96.02	
D: Aggregate-high p.type	97.03	
D: Aggregate-low p.type	94.96	
Ma: Aggregate-both p.types	97.77	
Ma: Aggregate-high p.type	98.42	
Ma: Aggregate-low p.type	97.13	
Mp: Aggregate-both p.types	97.87	
Mp: Aggregate-high p.type	98.66	
Mp: Aggregate-low p.type	97.10	
D: Proposer-high p.type	61.15	65.93
D: Responder-low r.type	35.88	34.07
D: Proposer-low p.type	45.22	45.49
D: Responder-high r.type	49.74	54.51
Ma: Proposer-high p.type	59.40	65.93
Ma: Responder-low r.type	39.02	34.07
Ma: Proposer-low p.type	42.32	45.49
Ma: Responder-high r.type	54.80	54.51
Mp: Proposer-high p.type	65.77	70.44
Mp: Responder-low r.type	32.89	29.56
Mp: Proposer-low p.type	35.10	34.96
Mp: Responder-high r.type	62.00	65.04

Efficiency data are based on accepted proposals. Aggregate efficiency is measured as the ratio of the sum of the proposer's and responder's actual payoffs to the same sum of payoffs that would have been obtained at Pareto optimum. Individual (proposer/responder) efficiency is the ratio of individual proposer's or responder's actual payoff to the same sum of payoffs for both players at the Pareto optimum. Note that aggregate and individual efficiency measures use the same denominator, namely the Pareto optimal payoffs to both players, hence the sum of proposer and responder efficiencies in matched cells gives back the aggregate efficiency in the corresponding cell (e.g., $61.15 + 35.88 = 97.03$). The final column shows the hypothetical decomposition of aggregate efficiency at Pareto optimum into individual ones, assuming that proposers take all private points. p.type=proposer type, r.type=responder type. PO=Pareto optimum

As Table 10 reveals, *aggregate* efficiency is very high across all treatments in excess of 90%. The rows for each treatment (D, Ma, Mp) labeled “Aggregate-both p.types” repeat the efficiency measures reported on earlier in Table 2. These

Footnote 14 (continued)

reported in Table 10. On the other hand, in matches between low proposers and high responders, there is more ambiguity as to which type fares better in terms of relative payoffs.

Table 11 Nonparametric tests for differences in efficiency

	Across treatment (SUM)			Pro. vs Res. (INDIV)	
	Alt.H	<i>p</i> -values		Alt.H	<i>p</i> -values
High	D<Ma	0.087	High Pro	D: P>R	0.022
P.Type	D<Mp	0.038	vs	Ma: P>R	0.022
	Ma≠Mp	0.602	Low Res	Mp: P>R	0.022
Low	D<Ma	0.125		Pool: P>R	0.0004
P.Type	D<Mp	0.300	Low Pro	D: P<R	0.112
	Ma≠Mp	0.753	vs	Ma: P<R	0.022
Both	D<Ma	0.125	High Res	Mp: P<R	0.022
P.Types	D<Mp	0.071		Pool: P<R	0.0006
	Ma≠Mp	0.917			

SUM=Mann–Whitney tests of differences in aggregate efficiency between two different treatments; INDIV=Wilcoxon signed rank tests of differences in individual efficiency between high-type proposer and low-type responder, or vice versa, within each treatment or in all 3 treatments pooled together. The session level data for aggregate and individual efficiency are based on accepted proposals. The columns ‘Alt.H.’ state the alternative hypotheses that aggregate efficiency are *higher* in D than in Ma or Mp (1-sided) or aggregate efficiency are *not* the same between Ma and Mp (2-sided) in the left part; and that individual efficiency is *higher* for high-type proposer/responder than for low-type responder/proposer within each treatment or all treatments pooled together (1-sided) in the right part. D=Discretionary, Ma=Mandatory-aligned, Mp=Mandatory-polarized. P.Type=proposer type, Pro.=proposer, Res.=responder

aggregate numbers are recalculated according to whether the proposer was a high or low type. The final 12 rows of Table 10 show proposers’ or responders’ average *share* of the Pareto optimal payoff and these percentages are further distinguished by the player’s type, high or low. We further report the Pareto Optimum (PO) share predictions for comparison purposes. Note that the actual efficiency shares for the Proposer and the Responder add up to the aggregate actual efficiency percentage in the first part of the table. For example, the numbers in the row “D: Proposer-high p.type” and in the row “D: Responder-low r.type” sum up to the numbers in the row “D: Aggregate-high p.type.”

As Table 10 reveals, we observe slightly higher aggregate efficiency in the mandatory treatments relative to the discretionary treatment, but as the Mann–Whitney test results using session level averages in Table 11 reveal, these differences are only significant when the proposer is a high type. This finding is surprising since the mandatory treatments should lead to higher efficiency due to the role played by the endogenous status quo default. However, as we have already noted, there is over-allocation to the public good in the discretionary treatment and this behavior raises payoffs to levels that are not far from the Pareto optimal benchmark. At the same time, in the two mandatory treatments, the desire for equal sharing is more pronounced (see the finite mixture model results and the allocations in the 2D simplex) and these fairness concerns reduce allocations to the public good. The net effect of

these two behaviors is to move payoffs in all treatments to be closer together so that efficiency differences across treatments are minimal. Further, we do not find efficiency differences between the two mandatory treatments.

While there are no large differences in aggregate efficiency across treatments, Tables 10 and 11 reveal an interesting difference in individual efficiency by player type. Specifically, we find strong evidence that high type proposers or high type responders achieve a greater share of total payoffs (greater efficiency) than do (their matched) low type responders or proposers. The differences are consistent with what would be predicted in the Pareto optimum (PO) as also reported in Table 10 (under PO predictions), and mainly reflect the fact that high types get more utility value from public good allocations than do low types. We summarize these findings as follows:

Result 7 Efficiency (actual payoffs achieved relative to the Pareto optimum) is high across all treatments in excess of 90%. The evidence for Hypothesis 6 is mixed with aggregate efficiency being marginally significantly higher in the Mp treatment but not in the Ma treatment as compared with the D treatment. Further, high types achieve a significantly larger individual share of the efficient aggregate payoff level regardless of whether they are in the proposer or the responder role.

5 Conclusions and suggestions for future research

We have reported on an experimental test of a model of public good bargaining due to Bowen et al. (2014). The main innovation of this model is the consideration of mandatory versus discretionary bargaining rules for public good provision. Under mandatory rules, the status quo level of public good provision becomes endogenous; once parties agree on a public good provision level, that level becomes the new status quo level. Thus, in the event of a break-down in bargaining between the two political parties, public good provision defaults to the status quo level which may be positive unlike in the discretionary case where the status quo level or the disagreement value is always zero. Theoretically, the problem of underprovision of the public good in the discretionary environment can be eliminated in the mandatory setting because the mandatory bargaining rules raise the bargaining power of the out-of-power party. Indeed, under mandatory rules, efficient public good provision becomes possible. The aim of our experiment is to test this important insight.

We consider both discretionary and mandatory bargaining rules and in the latter case, we further consider the degree of political polarization of the two parties as measured by differences in the weights that they attach to public good provision.

Consistent with the theory, we find that public good allocations are significantly higher under mandatory budget rules than under discretionary rules and that under the mandatory rules, pairs of players are very close to achieving the efficient level of public good provision. Still, they fall just short. What can explain this behavior? As we have seen, acceptance rates are increasing in both the public good amount and the private points offered to responders. The latter result is inconsistent with the theoretical prediction that proposers exercise full proposer power, but it is consistent

with findings from many ultimatum bargaining experiments. At the same time, low type proposers are not offering as large an allocation to the public good as high type proposers are (while the former types over-allocate and the latter types under-allocate on average, with respect to the equilibrium predictions conditional on the realized status quo levels for the public goods, in the mandatory treatments) and this behavior by the low types largely accounts for the shortfall in public good provision relative to the Pareto efficiency benchmark. This bias by the low types may reflect the low type's smaller payoff from the public good relative to the high types in combination with fairness concerns. Further, many proposers in the mandatory treatment, particularly the high types are choosing to reduce their own private points from equilibrium levels to fund the private points allocated to the responder. That is, proposers are heterogeneous in their exercise of proposer power, consistent with prior legislative bargaining experiments. The main difference of course, is that we are considering bargaining over public good provision which benefits all players. Here we observe that while our subjects fall short of achieving the efficient outcome, they do come tantalizingly close to reaching that benchmark.

Still, consistent with theoretical predictions, we find that as political polarization increases, both proposer types increase their allocations to the public good under the mandatory rules since the change in polarization leads to a higher efficient public good level. By contrast, under the discretionary rules, each proposer type makes public good allocations that are higher than the predicted static equilibrium levels, but that are further away from the Pareto efficient levels than under the mandatory rules. A main takeaway from our findings is that they help to rationalize the use of mandatory rather than discretionary budget rules in bargaining over public good expenditures between political parties.

We see several directions for future research on this topic. First, it would be useful to consider longer indefinite sequence lengths than in our study as that would allow more time for the status quo bargaining mechanism in the mandatory treatments to enable subject to possibly achieve convergence to the Pareto optimum. This change might be achieved by increasing the discount factor or by using the block random termination method of Fréchette and Yuksel (2017). Alternatively, we could consider changing the initial status quo level for the public good, e.g., to be at the Pareto optimum level. Second, it would be of interest to give subjects some pilot experience with several indefinite sequences involving both discretionary and mandatory bargaining rules and then ask them to choose which set of bargaining rules they would like to operate under—Bowen et al. (2017) suggest an interesting theory along this line. Third, it would be of interest to vary the duration of proposer power; we currently only consider a single value for p , the probability that a proposer remains in power. Changing p can affect the insurance motivation without changing the Pareto optimum public good levels, which is different from changing the degree of polarization ($\theta_H - \theta_L$). Finally, it would be of interest to connect our experimental design more closely with the Baron-Ferejohn legislative bargaining experiments, for example, by Fréchette et al. (2012) that involve three or more parties as well as both public and private goods in a dynamic bargaining game thereby enabling the study of majority rule rather than unanimous consent for implementation of bargaining outcomes. We leave all of these interesting extensions to future research.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10683-023-09818-6>.

References

- Baranski, A., & Kagel, J. H. (2015). Communication in legislative bargaining. *Journal of the Economic Science Association*, 1, 59–71.
- Baranski, A., & Morton, R. (2022). The determinants of multilateral bargaining: A comprehensive analysis of Baron and Ferejohn majoritarian bargaining experiments. *Experimental Economics*, 25(4), 1079–1108.
- Baron, D. P., & Ferejohn, J. A. (1989). Bargaining in legislatures. *American Political Science Review*, 83(4), 1181–1206.
- Battaglini, M., Nunnari, S., & Palfrey, T. R. (2012). Legislative bargaining and the dynamics of public investment. *American Political Science Review*, 106(2), 407–429.
- Battaglini, M., & Palfrey, T. R. (2012). The dynamics of distributive politics. *Economic Theory*, 49, 739–777.
- Bowen, T. R., Chen, Y., & Eraslan, H. (2014). Mandatory versus discretionary spending: The status quo effect. *American Economic Review*, 104(10), 2941–2974.
- Bowen, T. R., Chen, Y., Eraslan, H., & Zápal, J. (2017). Efficiency of flexible budgetary institutions. *Journal of Economic Theory*, 167, 148–176.
- Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9, 88–97.
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114(3), 817–868.
- Fréchette, G. R., Kagel, J. H., & Lehrer, S. F. (2003). Bargaining in legislatures: An experimental investigation of open versus closed amendment rules. *American Political Science Review*, 97(2), 221–232.
- Fréchette, G. R., Kagel, J. H., & Morelli, M. (2005a). Behavioral identification in coalitional bargaining: An experimental analysis of demand bargaining and alternating offers. *Econometrica*, 73(6), 1893–1937.
- Fréchette, G. R., Kagel, J. H., & Morelli, M. (2005b). Gamson’s law versus non-cooperative bargaining theory. *Games and Economic Behavior*, 51(2), 365–390.
- Fréchette, G. R., Kagel, J. H., & Morelli, M. (2005c). Nominal bargaining power, selection protocol, and discounting in legislative bargaining. *Journal of Public Economics*, 89(8), 1497–1517.
- Fréchette, G. R., Kagel, J. H., & Morelli, M. (2012). Pork versus public goods: An experimental study of public good provision within a legislative bargaining framework. *Economic Theory*, 49(3), 779–800.
- Fréchette, G. R., & Yuksel, S. (2017). Infinitely repeated games in the laboratory: Four perspectives on discounting and random termination. *Experimental Economics*, 20, 279–308.
- Moffatt, P. G. (2016). *Experimentics: Econometrics for experimental economics*. Macmillan: Red Globe Press.
- Palfrey, T. R. (2015). Experiments in political economy. In: Kagel, J. H. and Roth, A. E., editors. *Handbook of Experimental Economics*, 2, 347–434.
- Sherstyuk, K., Tarui, N., & Saijo, T. (2013). Payment schemes in infinite-horizon experimental games. *Experimental Economics*, 16, 125–153.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

John Duffy^{1,2} · SunTak Kim³ 

✉ SunTak Kim
yhvh0826@gmail.com

John Duffy
duffy@uci.edu

¹ Department of Economics, University of California, Irvine, USA

² ISER Osaka University, Osaka, Japan

³ Department of Economics, National Taiwan University, Taipei, Taiwan