

Sharing the Burden of Endogenous Negative Externalities*

Andrzej Baranski[†] Duk Gyoo Kim[‡]

November 5, 2023

Abstract

We study a game in which players negotiate the allocation of a negative externality, such as pollution-induced economic costs. Our goal is to explore the feasibility of preventing externalities through *ex-post* negotiations. We demonstrate that the unanimity rule results in complete pollution due to the veto power of players, allowing them to avoid paying more than their proportional share. Conversely, under the majority rule, multiple equilibria emerge. Pollution can be avoided if players are expected to form a coalition to penalize the largest polluter, thus establishing a credible threat of liability. However, experimental findings indicate the inefficacy of both rules in reducing pollution. Although a significant proportion of high polluters are held accountable, pollution persists due to instances where high polluters use their agenda-setting power to avoid paying. Our study underscores the muted influence of equity considerations when bargaining over costs in obtaining efficient outcomes, which has important implications for ongoing climate change loss and damage negotiations.

JEL Classification: C72, C92, Q52

Keywords: externalities, bargaining, experiment, burden sharing, inequity

*We thank Emin Karagozoglou for comments and suggestions. Baranski gratefully recognizes financial support by the Center of Behavioral Institutional Design and Tamkeen under the NYU Abu Dhabi Research Institute (Award CG005). Kim gratefully acknowledges financial support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)–Project-ID 139943784–SFB 884 and thanks Daeun Hwang for her research assistance.

[†]Division of Social Science and Center for Behavioral Institutional Design, NYU Abu Dhabi. a.baranski@nyu.edu

[‡]Department of Economics, Sungkyunkwan University, kim.dukgyoo@skku.edu

1 Introduction

Decades of sustained greenhouse gas emissions and other polluting activities have led to serious environmental degradation (Mora et al., 2018). An important question that has sparked heated debates within and between countries is how to share the burden associated with reparations, mitigation, and damages compensation (Colman and Mathiesen, 2022; Sengupta, 2021). The landmark agreement to create a fund for loss and damage resulting from climate change at the 27th Conference of Parties of the United Nations Convention on Climate Change on November 2022 will be tested in the upcoming Conference of Parties (COP28) when the countries negotiate their contributions to the relief fund. Similar contentious negotiations may arise when companies in a joint venture are liable for harms their products or services may have caused. Legal disputes to determine the share of the burden each firm will bear are common. A related problem arises after a bellic conflict when countries that fought together as allies discuss reparation funds for affected nations.

Although the aforementioned settings differ in important dimensions, they share five key characteristics. First, property rights are not fully defined, meaning it is not clear who is responsible and in what proportion should each party be held liable for the externality costs. Second, the sharing of the burden for reparations (i.e., the externality social cost) is decided *ex post*. Third, cost-sharing agreements are likely the result of negotiations between the parties involved, typically in a multilateral framework. Although these negotiations may occur under alternative consensus requirements, unanimous consent seems to be the norm absent an enforcing authority. Fourth, failing to reach agreements on how to share the costs can increase the magnitude of the externalities. Fifth, the notions of what constitutes a *fair share* are likely to differ between the parties, making it difficult to reach timely agreements.

In this article, we study a model of posterior bargaining over the cost sharing of reparations for endogenous externalities, which we refer to hereafter as pollution. Our aim is to understand if pollution can be mitigated under *ex post* bargaining and how different voting rules (majority vs. unanimity) and social costs (i.e., costs associated to the harms caused by pollution) affect burden-sharing. Similarly, we investigate the efficiency implications of each voting rule on the level of pollution and its mitigation. To do so, we first provide a set of hypotheses derived from a game-theoretic model, which we subsequently investigate by means of a laboratory experiment.¹ It is worth highlighting

¹Ideally, one would wish to have an empirical and externally valid measure of behavior, however, to the best of our knowledge, there are no natural experiments or data sets that would allow us to answer our research question. The theory and laboratory experiments presented here will serve as a first attempt at investigating cost sharing over endogenous externalities. Future research will help clarify the generality of the findings reported here.

that the existing literature has focused on the prevention of externalities through ex-ante contracting or bargaining, which we discuss in Section 2, while our focus is on ex-post negotiations.

In our model, players simultaneously and independently choose a production level, which generates immediate private benefits and pollution (social costs) that needs to be internalized ex post. Specifically, all players are perfectly informed of the preceding production and pollution decisions, and a player is selected at random to submit a proposal on how to share the pollution costs. The proposal is then observed by all and put up for vote. Depending on the agreement rule (majority or unanimity), if the required number of *yes* votes is obtained, the cost-sharing proposal is binding. Otherwise, the costs increase, and each member is responsible in equal parts for the increased² total costs. Our key question is whether there exists an equilibrium in the bargaining stage that can select efficient pollution levels. As we show, this depends on (1) the magnitude of the externality costs, (2) the voting rule, and (3) the way in which the externality costs are allocated.

Under unanimity, it is never possible to hold maximal polluters accountable because they will never accept to pay more than their outside option, unless they voluntarily decided to assume the costs. Therefore, when externality costs are below a certain threshold, full pollution occurs.³ We argue that the unanimity rule is closest to the setting of international relations and climate change negotiations, which we aim to model for two reasons. First, no country can be bound or forced by others at a Conference of Parties to pay for damages. In effect, each country can veto. Second, for a coalition to force another country to pay, they must do so by threat of force or commercial retaliation (i.e., imposing tariffs or quotas; see Nordhaus (2015) for a proposal in this direction). Commercial retaliation will certainly trigger disputes in the World Trade Organization and the threat of military actions has not been alluded as possible means to enforce *cooperation* on climate change mitigation and reparation matters. Hence, we believe that the unanimity voting rule captures the practical difficulties in reaching agreements on loss and damage transfers.

Under the majority rule, parties can form coalitions to penalize the highest polluters. Thus, there exist equilibrium negotiation strategies that credibly reduce the incentive to pollute. However, there are also equilibria that do not lead to pollution deterrence, for

²The increases in the magnitude of the costs may have two non-exclusive interpretations. First, reaching an agreement to deal with the effects of pollution may be less costly compared to each country dealing with the problem on its own (i.e., disagreement). Second, the magnitude of the damage caused by pollution may increase if not addressed.

³Of course, if pollution costs are prohibitively high, expected burden makes polluting unattractive because it is not individually rational to do so. The problem is only interesting when it becomes a social dilemma in the parameter region we study.

example, when high polluters are not singled out or penalized. Our experimental investigation focuses on understanding if and when such a pollution-detering equilibrium is played. We view the majority rule case as a hypothetical scenario, where both theory and experiments can shed light on plausible behavior should these conditions arise in the *real-world* settings of interest.

Our model bears a close resemblance to settings where, instead of sharing costs, the sharing of profits resulting from joint production is being decided through bargaining (Gantner et al., 2001; Cappelen et al., 2007; Rodriguez-Lara and Moreno-Garrido, 2012; Gantner et al., 2016; Baranski, 2016). We clarify that these problems are not isomorphic, and hence our setup is not simply a reframing of a well-studied case. Specifically, in a game of endogenous profit distribution with the same timing and structure as ours, there would be a unique bargaining equilibrium which entails giving each player their outside option (0, in our case), regardless of the voting rule.⁴ Importantly, the experimental literature on bargaining to distribute an endogenously produced surplus (which we review in detail in Section 2) provides unequivocal evidence that entitlements and contributions matter to bargainers when deciding how to allocate benefits. Higher contributors typically receive larger shares. In this sense, we believe that preferences for equitable sharing are likely to shape bargaining behavior over endogenous social costs too. Thus, even if our theory posits that ex-post bargaining under unanimity is not useful in deterring externalities, the empirical evidence suggests that it might in the laboratory if subjects abide by equitable sharing of the costs.

Our experimental results show that bargaining over externality cost sharing under the unanimity rule fails to prevent pollution. This is mainly because every player has veto power over any proposal that allocates more costs than private benefits, and as such, no proposals can enforce a cost large enough to prevent pollution. The findings under unanimity are consistent with our theoretical predictions.

Under a majority rule, we find that most proposals impute the total costs to a single person, but this person is not always the highest polluter. When the cost of pollution is low, we find no difference between unanimity and majority in the amount of polluting activity (as predicted by our theory). However, for a high cost of pollution, we find that there is a mild moderation of pollution, especially as subjects gain experience in the game. This is consistent with the moderating effect that the threat of being assigned a high share of the costs can have on incentives to pollute, but the effect is nowhere close to full deterrence. Although we do find evidence of targeting high polluters to pay for the externalities, there is also a substantial portion of outcomes in which low polluters are

⁴Because the proposer can exploit her bargaining power, no one would have an incentive to contribute for the production of joint profits. Therefore, there is no equilibrium where full efficiency arises.

paying a high fraction of the costs while high polluters are spared. The presence of such inequitable allocations, we argue, mutes the deterrent incentives.

Thus, in a nutshell, we conclude that ex-post burden-sharing negotiations under both voting rules do not aid in mitigating pollution significantly. This result shows that the equitable sharing norms typically observed in bargaining games with joint production do not carry over with the same intensity to case of accountability for jointly produced social costs.

Efficiency in our setting is affected not only by the externality costs, but also by the bargaining behavior. As mentioned earlier, when groups fail to reach an agreement, the externality becomes larger. We find that as the externality cost increases (a treatment variable), the rate of agreements in bargaining decreases (under both rules). Thus, we uncover a novel efficiency-affecting factor in negotiations: When the endogenous cost to be shared is larger, gridlock is more prevalent, further affecting efficiency. This trade-off is absent in theory, where disagreements should never occur in equilibrium. Beyond the theory, we are unaware of any other bargaining experiment that investigates how the size of the costs to share affect efficiency.

Our research contributes to understanding how bargaining can help prevent externalities. Previous work by [Pigou \(1932\)](#) and [Coase \(1960\)](#), has focused on dealing with the internalization of externalities ex ante, that is, before polluting decisions are made. We offer a different framework, which we have argued resembles the timing of actions in several settings of interest. Experimental research on the resolution of collective dilemmas with applications to climate change and environmental conservation is vast and growing ([Ostrom et al., 1994](#); [Rodriguez-Sickert et al., 2008](#); [Stranlund et al., 2011](#); [Hauser et al., 2014](#); [Ghidoni et al., 2017](#); [Calzolari et al., 2018](#); [Pevnitskaya and Ryvkin, 2022](#); [Alberti and Mantilla, 2023](#)). We address several of the most closely related works, especially those dealing with the prevention of externalities in [Section 2](#).

The article proceeds as follows. In [Section 2](#) we review the related literature. In [Section 3](#) we present the model and equilibrium predictions. [Section 4](#) contains the experimental design and [Section 5](#) the hypotheses to be tested. We conclude in [Section 7](#).

2 Related Literature

The problem we study here relates to several strands of the literature. In this section, we discuss the relationship to the early literature on the handling of externalities by Arthur Pigou and Ronald Coase. Next, we draw parallels with experimental investigations on the problem of externalities and public good provision, and the formation of

coalitions to fund public goods. Finally, we relate our game to the literature on bargaining over an endogenous positive fund.

We contribute to understanding mechanisms for the mitigation of externalities and the production of public bads. A Pigouvian tax (Pigou, 1932) to induce internalization of social costs would require in our setting that each party *pays* for her share of the pollution-induced social costs. In principle, an outcome that is payoff equivalent to a setting with Pigouvian taxation could be achieved in a *fair* negotiation if the parties are willing to propose and accept such burden-sharing schemes *ex post*. However, because the externalities are sunk at the moment of bargaining, a self-regarding and rational individual would never accept to pay more than what her outside option requires her. Moreover, a rational bargainer would never claim a larger share of the costs when she can strong-arm others to pay more than her. Thus, our experimental investigation can shed light into the nature of bargaining behavior under endogenous social costs and whether proportionality in cost-sharing obtains or strategic self-regarding behavior arises.

Our work is clearly related to the Coasean bargaining approach (Coase, 1960) to deal with externalities, but there are two key differences. First, in the Coasean context, one party benefits from the right to pollute, while the other benefits from its absence. The first difference is that in our setting, all parties benefit from *polluting* and all benefit when others are held responsible for associated losses and damages. Second, Coase argues that assigning property rights will allow parties to bargain over compensation *ex ante*, that is, parties will internalize the costs of externality with certainty. In our setting, bargaining occurs *ex post*, and property rights are technically undefined: Anyone has the same right to propose a burden-sharing agreement. Although we do not argue that one approach is *better* than the other, we believe that our setting captures more accurately the current global discussions of environmental damage reparations and the anarchic geopolitical framework with no *de facto* property rights at the negotiation stage.⁵

There are three experimental investigations concerning the internalization of externalities that are closest to ours. First, Dekel et al. (2017) study a setting in which the provision of a public good is efficient (maximizing the aggregate reward) but harms a minority of society's members. Subjects play a linear public goods game in which contributions to the public good increase the sum of payoffs, but private returns are lower than the private cost of contribution, resulting in no contribution in equilibrium. Dekel et al. find that allowing for ex-post voluntary rewards can lead to efficient provision, but this happens only when players can communicate prior to playing the game. The reason

⁵The language adopted in the COP27 refers to the creation of a fund for "loss and damage" related to climate change, which do not imply liability in legal terms. The United States has signaled that it will not accept any agreement that involves liability (Sengupta, 2021). In terms of our model, we interpret the situation as if there are no property rights about who deserves compensation and by whom.

is that those who are harmed by the public good seek compensation, which is commonly offered by those who benefit from the provision. Second, in a similar setting where some parties are affected by the public good, [Alberti and Mantilla \(2023\)](#) experimentally investigates a mechanism by [Van Essen and Walker \(2017\)](#) that allows the compensation of affected parties. In their game, provision is only possible under unanimous agreement, and players can negotiate ex-ante transfers to compensate those harmed by the provision of the public good. As in [Dekel et al. \(2017\)](#), the authors find that communication channels are important in obtaining efficient outcomes by coordinating appropriate transfers to compensate those injured. The third related work is [Pevnitskaya and Ryvkin \(2022\)](#), who study a dynamic game of public bad production and endogenous abatement. In their experimental game, subjects can invest in a clean technology to reduce their pollution. They find that when such technology is of public access (i.e., if one player invests in it, others benefit) pollution is the lowest. Furthermore, in line with [Alberti and Mantilla \(2023\)](#), they report that communication aid in achieving more efficient outcomes.⁶

Our work is also related to the study of how public goods are provided through collective decision-making. [Hamman et al. \(2011\)](#) experimentally study a public good provision game in which players can delegate contribution decisions to an elected member by majority vote. The authors report that, relative to the decentralized standard public goods game, elected decision makers typically select the most efficient outcome and distribute the burden of the public good provision relatively evenly among players. In this game, elected decision makers have the option of expropriating the minority by imposing the burden on a subset of players (which would be in line with standard equilibrium predictions), but this happens very rarely in the laboratory. Similarly, in our game, the player selected to propose may submit a *fair burden-sharing plan* or can exploit a minority strategically and avoid paying her share of the social costs.

The ex-post burden sharing stage of our game shares similarities with the widely-studied punishment mechanism in public goods introduced by [Fehr and Gächter \(2000\)](#). The punishment mechanism allows players to individually deduct payoffs from other players at a personal cost, which serves to discipline would-be free-riders. In our setting, the cost-sharing scheme may serve the same purpose as the punishment mechanism, except that it needs to be agreed upon by the required quota (majority or unanimity), and the punishment can be executed at zero personal costs. Furthermore, the available points to deduct are endogenous in our game, as these depend on the level of externalities produced by each player.

The game we study is also related to the common pool resource problems described

⁶See [Calzolari et al. \(2018\)](#) for an experiment with persistent accumulation of pollution in a dynamic setting. See [Ghidoni et al. \(2017\)](#) for an experiment on an indefinite repetition with non-cumulative externalities.

in [Ostrom et al. \(1994\)](#) and [Ostrom \(2006\)](#). In these settings, a common pool resource can be extracted for personal gain (e.g., water from irrigation canals, wood, and clean air), but extraction creates a personal and social cost. These costs are not fully internalized, and because exclusion from consumption is difficult, over-extraction occurs. In our game, subjects' decisions are framed as extractions, which create a social cost. However, the sharing of the social costs is not predetermined; instead, bargaining determines the split. This feature distinguishes our setting from existing studies where the assumed technology or mechanism defines both extraction payoffs and distribution of the social costs. For example, [Rodriguez-Sickert et al. \(2008\)](#) allow for the endogenous choice of a sanctioning mechanism that fines extractors. Thus, when the sanctioning technology is endogenously selected, this can affect the final payoffs. The authors found that this variant helps to foster efficiency, in line with their theoretical benchmark. [Hauser et al. \(2014\)](#) report on a dynamic extraction experiment in which a future generation (i.e., another set of subjects) suffers the consequences of overextraction. They find that voting on a binding extraction level leads to efficient extraction. [Abatayo and Lynham \(2016\)](#) investigate the role of endogenous monitoring and communication, and find that communication decreases extraction.

Finally, we speak to a broad literature on bargaining over an endogenous fund which has, to the best of our knowledge, exclusively focused on the distribution of benefits and not endogenous costs, as we do. Subjects in experiments display preference for distributing the jointly produced fund in a manner that reflects contributions. For example, [Cherry et al. \(2002\)](#) report lower transfer in a dictator game where the dictator's fund was determined by performance in a quiz. Similarly, [Gantner et al. \(2001\)](#) find evidence of equitable sharing (i.e., proportionality) in an ultimatum game where both subjects in a dyad were able to have the opportunity to invest in a joint project, and the profits would be subsequently bargained over. [Stoddard et al. \(2014\)](#) study behavior in a public goods game in which the total fund produced by voluntary contributions is split by a randomly-selected allocator, and report that equitable sharing promotes efficiency. Beyond the lab, [Van Dolder et al. \(2015\)](#) report that participants in a TV show who earned money as a team also display preferences for sharing proportionally to individual contributions. Thus, equitable sharing ([Adams, 1963](#)) is a strong driver of behavior in bargaining when the origin of the fund to split is endogenous.

3 The Model

3.1 Description of the Game

In this section, we present a simple model of bargaining over endogenous costs. Consider a three-person⁷, two-stage game where the first stage involves the extraction of private benefits which create social costs, and the distribution of these costs is determined in the second stage. Three players are indexed by $i \in \{1, 2, 3\} \equiv N$. In the first stage, every player is endowed with $E > 0$ units of resource and player i claims $g_i \in [0, E]$ for her benefit. The total sum of claims generates aggregate costs of $C = \alpha \sum_i g_i$, $\alpha \in (0, 3)$.⁸ One may regard that the claimed amount corresponds to the activities beneficial to self but harmful to society, such as profitable productions that produce pollution, and in this respect, α is a parameter that describes how bad the production technology is for the environment. In the following, we refer to g_i as *pollution* (for the sake of private benefits) and C as the aggregate *costs* induced by the total sum of pollution.

In the second stage, players collectively decide how to allocate the costs C . The vector of claims is public information. Specifically, the collective decision is made in the following way: One of the three players is randomly selected, with equal probability, to propose a split of the costs. Formally, we denote a proposal by $p \in \mathcal{P}$, where $\mathcal{P} = \{(p_1, p_2, p_3) \in [0, 1]^3 \mid \sum_i p_i = 1\}$. In words, p describes the proportion of the costs that each player is charged. Then every player votes for or against the proposal. If $q \in \{2, 3\}$ or more players vote for the proposal, it is approved, and player i earns the payoff of $g_i - p_i C$. If the proposal is rejected, player i accrues the payoff of $g_i - \frac{C}{2}$. Our model assumes that the magnitude of the costs increases upon disagreement.⁹ Such increases in the disagreement costs may have two non-exclusive interpretations. First, reaching an agreement to collectively deal with the social costs may be more effective than each player (i.e., country in the context of pollution) dealing individually with the social costs. Second, the magnitude of the social costs caused by pollution may increase if not dealt with in a timely manner.¹⁰ We further assume that the players' utility is linear in their payoffs and depends only on them.

⁷All analyses in this section are valid for $n \geq 3$ odd players. We focus only on the case of three players for the close link between the model and the experiment.

⁸We consider $\alpha < 3$ because otherwise, it is trivial that taking positive claims with expecting at least a fair share of the costs is strictly dominated by zero claims.

⁹The aggregate payoff for bargaining agreement is $\sum_i g_i - \sum_i p_i C = \sum_i g_i - C$ and for disagreement is $\sum_i g_i - \frac{3}{2}C$.

¹⁰Related to the second interpretation, if the magnitude for the social costs upon disagreement were to be smaller than that upon agreement, it would directly imply that any actions intended to discourage agreement, for example, sweeping the impending problems under the rug or sabotaging the bargaining process, are optimal. Smaller disagreement costs may also mean that the social costs in the discussed agenda are unimportant so that those do not require immediate attention and timely actions.

3.2 Equilibrium Characterization

Each player's strategy consists of the amount of claims, the proposal when selected as a proposer, and the voting decision when not selected as a proposer. As typically assumed in the literature, we assume that whenever a player is indifferent between voting for and against the proposal, she will vote for it. The subgame-perfect Nash equilibrium (SPE) is our solution concept. If $q = 3$, the SPE is essentially unique, but otherwise this game has a continuum of subgame-perfect equilibria (SPEa). Before describing equilibria for all cases, it is worth mentioning that non-proposer's second-stage equilibrium strategy is straightforward: Player i votes for the proposal if and only if the costs assigned to her are not greater than the costs she would have when the proposal is rejected.

Lemma 1. *In any equilibrium, player i 's optimal voting behavior is to vote for the proposal if and only if $p_i \leq \frac{1}{2}$.*

With Lemma 1 in mind, it is straightforward that any equilibrium proposal involves the smallest costs to the proposer herself. In our parametric setting where the outside option is $\frac{C}{2}$, the proposer can impute the totality of costs among the voters, assigning 0 to herself. Lemma 2 states this observation.

Lemma 2. *Let player i denote the randomly selected proposer. In any equilibrium, $p_i = 0$.*

Accordingly, the bargaining behavior under unanimity is simple: All players claim their private benefits fully, that is, pollute as much as they can, and whoever becomes the proposer assigns half of the entire costs to two other players.¹¹

Proposition 1. *When $q = 3$, the following strategy profile is the essentially unique SPE: (1) For all i , $g_i^* = E$, (2) proposer i offers $p_i^* = 0$ and $p_j^* = \frac{1}{2}$ for $j \neq i$, and (3) player j votes for the proposal if $p_j^* \leq \frac{1}{2}$. The proposal is approved in this equilibrium, as all players vote for it.*

In the following proposition, we characterize the set of SPE proposals under the majority rule.

Proposition 2. *Let $q = 2$ and i denote the proposing player. The SPE strategies in the bargaining subgame are as follows. (1) The proposer assigns $p_i^* = 0$ to herself; (2) The proposer distributes the entirety of the costs between both other voters; (3) player j votes for the proposal if $p_j^* \leq \frac{1}{2}$.*

¹¹Recall that $\alpha \in (0, 3)$. If $\alpha > 3$, no player would want to pollute because it creates an expected cost larger than the benefits.

It is easy to see that any such proposal will be voted in favor by at least one other nonproposer and thus be approved. We now show that pollution decisions will depend on the equilibrium expected to be played out in the bargaining subgame. We first describe an equilibrium in which the proposer selects one member at random to pay for the entire costs.

Proposition 3 (Random Cost Allocation). *When $q = 2$, the following strategy profile is another SPE: (1) For all i , $g_i^* = E$, (2) proposer i randomly selects a non-proposer $k \neq i$ with equal probability and proposes $p_k^* = 1$ and $p_{-k}^* = 0$, and (3) player j votes for the proposal if $p_j^* \leq \frac{1}{2}$. The proposal is approved in this equilibrium, as two players vote for it.*

The intuition for these results is straightforward. If being assigned the costs occurs at random, every player has 1/3 chance of paying for pollution. This induces an expected cost of $\alpha/3$, which is less than the benefit of polluting. Thus, to be able to deter pollution, there must exist an expectation of being assigned a share of the costs that makes it unprofitable to pollute. In the following proposition, we describe one such possibility.

Proposition 4 (Allocation to the highest). *When $q = 2$ and $\alpha > \frac{3}{2}$, the following strategy profile is another SPE: (1) For all i , $g_i^* = 0$, (2) proposer i picks player $k \neq i$ whose $g_k = \max_{j \in N \setminus i} g_j$, proposes $p_k^* = 1$ and $p_{-k}^* = 0$, and (3) player j votes for the proposal if $p_j^* \leq \frac{1}{2}$. The proposal is approved in this equilibrium, as all players vote for it.*

To see why, suppose that all players are claiming the same positive amount. Then, each player expects to pay the costs with probability 1/3. However, under these bargaining strategies, one player can undercut and pollute marginally less (say, ϵ). Then, she forgoes ϵ in payoffs, but avoids being assigned the totality of costs with certainty, which increases her payoff. A similar logic can be used to see why there is no asymmetric vector of claims.

Note that the equilibrium in which the total costs are allocated to the highest polluter creates the strongest deterrence. Any other rule attaching a lower probability of proportion of costs to the highest polluter would necessarily dilute the expected costs from polluting. Hence, if preventing pollution under the equilibrium described in Proposition 3 is not possible, it is also not possible under any other allocation of costs.

4 Experimental Design

We consider six (three by two) different treatments which vary in two dimensions: the magnitude of public bads production ($\alpha \in \{0.8, 1.2, 1.6\}$) and the voting rule (majority

or unanimity). The six treatments are abbreviated as M08 (a majority rule with $\alpha = 0.8$), M12 (majority with $\alpha = 1.2$), M16 (majority with $\alpha = 1.6$), U08 (a unanimity rule with $\alpha = 0.8$), U12 (unanimity with $\alpha = 1.2$), and U16 (unanimity with $\alpha = 1.6$). When necessary, we collectively call M08, M12, and M16 as the Majority treatments and the other three as the Unanimity treatments. We also collectively call M08 and U08 as the low cost treatments and M16 and U16 as the high cost treatments. Table 1 summarizes our experimental design.

Treatment	Voting Rule	Cost Multiplier	#Sessions	# Subjects
U08	Unanimity	0.8	2	24
U08	Majority	0.8	2	24
U12	Unanimity	1.2	3	36
M12	Majority	1.2	3	36
U16	Unanimity	1.6	3	36
M16	Majority	1.6	3	36

Table 1: Experimental Design

Subjects are randomly and anonymously placed in groups of three members. Each subject is endowed with 1,000 tokens¹² (the currency units used in this experiment) in their private account and informed that there are 600 tokens in their group account. In the first stage, subjects simultaneously and independently decide how many tokens to claim from the group account, any integer up to 200 tokens per subject. The total sum of tokens claimed in the first stage multiplied by the externality factor (i.e., 0.8, 1.2, or 1.6 depending on treatment) determines the total costs to be divided in the second stage. Individual claims are publicly revealed to all members of the group at the bargaining stage.

The second stage follows a standard random-proposer ultimatum protocol. Each member of the group submits a proposal establishing how the costs will be split. One proposal is randomly selected and immediately voted on. The proposal is approved when a qualified number (2 in Majority treatments or 3 in the Unanimity treatments) of members agree. If the proposal is accepted, the costs are charged according to the proposal. If the proposal is rejected, half of the total costs are charged to every member.

The process, which consists of the claim and bargaining stages, is repeated 5 times. We refer to each repetition as a period. In each period, subjects are randomly reshuffled to form new groups of three members and reassigned ID numbers, and thus subjects cannot identify each other across periods. One of the periods is randomly selected at the

¹²Subjects in U16 and M16 are endowed with 1,100 tokens to guarantee the theoretical lower-bound of the payment to be greater than the minimum. Ex post, the average payment of M16 and U16 is not statistically different from that of the other treatments ($p=0.915$).

end of the experiment to count for payment. At the end of the experiment, participants completed a short survey questionnaire asking for their gender, age, number of recognizable friends in the same session, and risk preferences. Sample experimental instructions can be found in Appendix C.

We used an interactive online platform called LIONESS (Live Interactive Online Experimental Server Software, [Arechar et al., 2018](#)). A total of 192 subjects were recruited from the undergraduate and graduate student population of Seoul National University. We had three sessions for U12, M12, U16, and M16 each and two sessions for U08 and M08 each. After the subjects joined an online meeting and their registrations were verified, the experimenter asked them to turn off the webcam and renamed their displayed names to two alphabet letters they arbitrarily chose so that their identities, hence decisions, remain anonymous to the experimenter as well as other subjects. Subjects were asked to read the instructions displayed on their screens carefully and pass a comprehension quiz. The average payment per subject was 14,426 KRW (about 12 USD). The payments were made by online transfer after receiving the personal payment code generated at the end of the experiment. Each session lasted 50 minutes.

5 Hypotheses

In this section, we present testable hypotheses about our experimental data based on our theoretical predictions. It is worth noting that we do not regard the equilibrium predictions as a normative suggestion of play, but take them as a basis for our null hypotheses. The theory is especially important in guiding our expectations about outcomes for which we have no previous empirical evidence to leverage. As will become clear, in some cases we do have empirical evidence that helps formulate more nuanced hypotheses. In formulating the hypotheses, we refer to the claimed tokens as *pollution*.

The first hypothesis concerns subjects' pollution choices. In all treatments except M16, the unique equilibrium pollution level (as stated in Propositions 1, 2, and 3) is that everyone would choose $g_i^* = E$.

Some subjects may be concerned with overall efficiency ([Andreoni and Miller, 2002](#); [Engelmann and Strobel, 2004](#)), and thus we expect lower pollution as the cost multiplier increases (within each rule). When the cost multiplier is below 1, social efficiency and individual payoff-maximizing behavior coincide in that full pollution is optimal. On the contrary, when the cost multiplier is greater than 1, there is a tension between social efficiency and individual payoff-maximizing incentives.

Hypothesis 1. *Within each voting rule, the amount of pollution decreases as the cost multiplier increases.*

Our theoretical results posit that it is not possible to curb pollution when the cost multiplier is 0.8 and 1.2 under both voting rules. This theoretical prediction leads to the null treatment effect in the voting rule.

Hypothesis 2. *When the cost multiplier is 0.8 or 1.2, pollution under majority and unanimity rule is the same, holding the cost multiplier fixed.*

However, when the cost multiplier is 1.6, there exists an equilibrium in which pollution is completely deterred under the majority rule. This equilibrium builds upon subject's particular bargaining behavior. Our empirical expectation at the burden-sharing stage, based on the findings of bargaining experiments in which the fund to distribute is endogenous (Konow, 2000; Gantner et al., 2001; Cherry et al., 2002; Rodriguez-Lara and Moreno-Garrido, 2012; Dong et al., 2019) is that subjects will try, by and large, to split total costs in a way that reflects individual pollution decisions. In particular, the highest polluter is most likely to receive the largest share of costs under the majority rule, where such a proposal can pass. Proportionality in sharing costs or targeting of the highest polluters is less likely to be approved under unanimity, hence less likely to be proposed, too.

We posit our next hypotheses with these observations in mind:

Hypothesis 3. *Under the majority rule, proposals that assign the largest cost share to the highest polluter are modal and are more commonly observed than under the unanimity rule.*

Hypothesis 4. *If the likelihood of being assigned a large share to the higher polluter is high under majority, then we expect that pollution is lower in M16 than in U16.*

Whereas the first four hypotheses pertain to how subjects pollute and how they propose to distribute the costs, our last hypothesis regards the response to the proposal as a voter. Recall that our experimental game clearly establishes what subjects will earn when the proposal is rejected. It implies that there is no rational ground for bargaining disagreements: the proposer offers the acceptable costs to other members, and the responders accept an offer if accepting it renders larger payoffs than rejecting it. Therefore, all proposals should be accepted in all treatments.

However, it is well established that rejections occur in similar bilateral (Cochard et al., 2021) and multilateral bargaining games (Baranski and Morton, 2021), and that unanimity entails a higher likelihood of disagreement. What is less known is whether the magnitude of the costs to distribute and the cost multiplier have any effect on the ability to reach agreements. Thus, we posit our null based on the theoretical benchmark.

Hypothesis 5. *Within each voting rule, the disagreement rate is unaffected by the cost multiplier and the overall size of the externality costs to distribute.*

6 Results

In this section, we report the experimental findings in the same order in which we have presented our hypotheses. Hypothesis 1 posits that pollution decreases as the cost multiplier increases (within each voting rule).

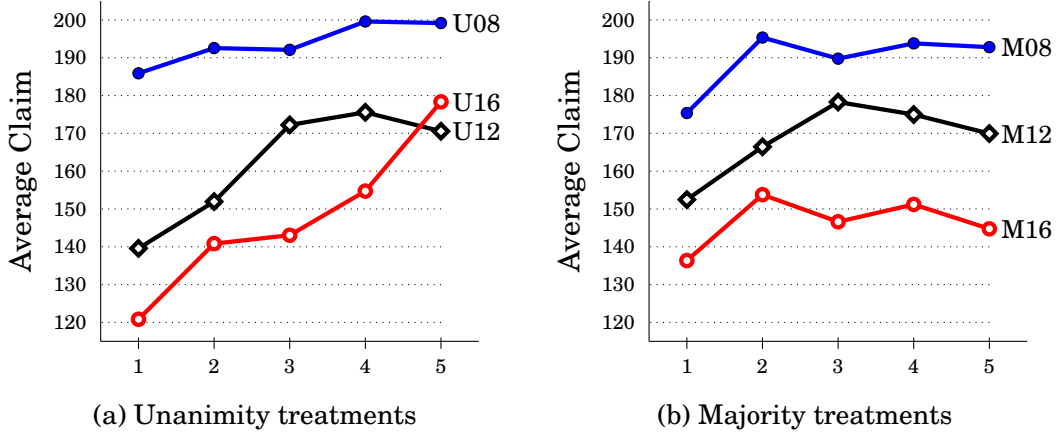


Figure 1: Average Claim by Period

Figure 1 shows the average pollution (claim of private benefits) by period. In the Majority treatments (Figure 1b), except for the slight increase after the first period, the average pollution is quite stable in level, and the higher cost multiplier is significantly associated with a lower average pollution in both Majority ($p < 0.001$) and Unanimity ($p < 0.001$) treatments¹³, which supports our first hypothesis that the concern for social efficiency when the cost multiplier is larger than 1 can curb pollution to some extent.

Result 1. *The amount of pollution decreases as the cost multiplier increases.*

We now turn to testing hypothesis 2, which posits that pollution levels under unanimity and majority rule are identical for the lower cost multipliers. Table 2 shows theoretical and observed levels of pollution as proportions of maximum pollution. When the cost multiplier is 0.8, the average level of pollution is close to 95% of the maximum pollution (94.70% in M08 and 96.93% in U08), and we observe no significant difference between the voting rule ($p = 0.229$). The average level of pollution is still high when the cost multiplier is 1.2 (84.22% in M12 and 80.99% in U12), and again, there is no significant difference in the voting rule ($p = 0.428$). These observations support our second hypothesis.

¹³Unless otherwise stated, we report the p of the estimated coefficient of the linear regression of the outcome variable on the control variable of interest, clustering standard errors at the individual level. Regression results with more control variables are in the Appendix B.

Treatment	Socially Optimal Pollution	Pollution in Equilibrium	Observed Pollution
U08	100%	100%	96.93%
M08	100%	100%	94.70%
U12	0%	100%	80.99%
M12	0%	100%	84.22%
U16	0%	100%	73.78%
M16	0%	0% or 100% [†]	73.27%

[†]: Prediction varies by equilibrium. See Section 3 for details.

Table 2: Theoretical and Observed Levels of Pollution
% as a proportion of maximum pollution

Result 2. *When the cost multiplier is 0.8 or 1.2, the average level of pollution under majority is not significantly different from that under unanimity, holding the cost multiplier fixed.*

We now turn to hypothesis 3, which states that proposals assigning the largest share of the costs to the highest polluter are modal under majority rule and more common compared to the unanimity rule. To investigate the nature of the proposals, we classify these into three main categories. We say a member is included in the distribution of costs if she receives a share that is at least 5% of the total costs to distribute. The three categories are: *one-ways splits*, where only one member is included, *two-ways splits* where two members are included and *three-ways splits* defined similarly.

We also consider two other types of splits, because as we will see, these are quite common. We say a proposal is *egalitarian* if the difference between minimum and maximum share of costs is less than 5% of the total costs. We refer to a proposal as *proportional* if the percentage of the proposed costs to each member and the individual’s contribution to the aggregate costs (individual’s claim times the cost multiplier) is within a 95%–105% range for all members.

Proposal Type		Three-way split	Two-way split	One-way split	Egalitarian	Proportional
Unanimity	All	0.852	0.123	0.025	0.273	0.313
	Accepted	0.981	0.019	0.000	0.433	0.413
Majority	All	0.433	0.007	0.496	0.075	0.110
	Accepted	0.400	0.034	0.566	0.062	0.103

Table 3: Types of the Submitted and Accepted Proposals

As shown in Table 3, proposals in the Unanimity treatments are quite different from

those in the Majority treatments.¹⁴ The egalitarian proposal is rarely observed in the Majority treatments, while it is common in the Unanimity treatments. It is also noticeable that proportional proposals are much more common under unanimity. Two-way split proposals are rare under both voting rules, although the common form of equilibrium in both voting rules involves allocation of the costs to two players (Propositions 1 and 2).

In the Majority treatments, most common proposals are one-way splits,¹⁵ which are slightly more common in M08 (60.0%) and M12 (47.8%) than in M16 (44.4%). It is worth noting that one-way splits are pervasive even in M08, where a proportional allocation can guarantee the positive payoffs to all three members. The majority of the case involve penalizing a member who claimed the most (77.8% in M08, 76.7% in M12, and 58.8% in M16), supporting our third hypothesis.

Result 3. *In the Majority treatments, proposals assigning almost all costs to one player are modal, but these are rarely observed in the Unanimity treatments. The recipient of the largest cost is typically the highest polluter.*

We now investigate why, despite the fact that high polluters often receive the largest share of the costs, we still observe high levels of pollution. For this purpose, we estimate a regression model to establish if there exists a correlation between that share of costs (as a proportion of total costs) a member is offered and her level of pollution (as a proportion of total pollution). We also control for whether the recipient is oneself, to test for the existence of a proposal-making advantage, or whether players making the proposals treat themselves more favorably. If the estimated coefficient for the relative pollution is equal to 1, this means that in expectations members are paying for the totality (or more) of the costs they created. If it is less than 1, members do not fully internalize the costs associated with their pollution decisions.

The results reported in Table 4 (in columns 1 and 3) reveal that there is a positive correlation between a player's pollution and her share of costs, but this is nowhere near a 1 to 1 relationship. The relationship appears stronger under the majority rule, implying that the egalitarian proposals largely observed under unanimity do not appear to be comprised of distributions of shares that reflect pollution choices strongly. We also find clear evidence that subjects assign themselves lower shares of the costs.

In the regression results reported in columns 2 and 4, we interact the relative pollution by a player with the dummy variable indicating whether the share is assigned to

¹⁴We mainly examine the proposal types of all the submitted proposals. Recall that every member submits a proposal, and only one of them is put up to the vote, the number of accepted proposals is much smaller than that of all submitted proposals.

¹⁵This observation is consistent with the findings of Kim and Lim (2019) who report that multilateral bargaining outcomes for the division of losses often involve an allocation to the entire losses to one person.

oneself. In the majority treatment, the estimate coefficient is -0.56 . The pollution coefficient is 0.56 . Hence, subjects propose splits that condition others' shares on pollution, but not their own.

Table 4: The determinants of Proportion of Costs Offered

	Majority		Unanimity	
	(1)	(2)	(3)	(4)
Pollution (relative)	0.38*** (0.08)	0.56*** (0.09)	0.15*** (0.05)	0.18** (0.07)
Share to self (0 or 1)	-0.16*** (0.02)	0.02 (0.05)	-0.06*** (0.01)	-0.04 (0.04)
Pollution×Share to self		-0.56*** (0.15)		-0.08 (0.12)
N	1440	1440	1440	1440
R^2	0.064	0.071	0.062	0.064

OLS regression of offered share of the costs. The unit of observation is a share of the costs offered by a subject to each member of the group. *Pollution* is the relative size of the claim in the group for a given recipient. *Share to self* is the indicator of whether the offered share is to herself. Standard errors clustered at the individual level of the subject making the offer are in parentheses. *, **, and *** indicate statistical significance at the 5% level, 1% level, and 0.1% level, respectively.

We now turn to our fourth hypothesis concerning pollution levels in M16 and U16. Recall that this is the only treatment where we have a theoretical possibility where equilibrium play (under standard assumptions) would yield a lower level of pollution. The results for M16 and U16 are similar to what we report in Table 4, so we relegate them to the Appendix B.

Result 4. *Although the likelihood of the largest polluter being punished is high under majority, high polluters often penalize others when proposing, and as such, the relationship between pollution and share of the costs is weak. As a result, the overall pollution in M16 is not significantly different from that in U16.*

We now turn to Hypothesis 5 which concerns the likelihood of reaching an agreement, namely, that within each voting rule, agreement rates are not affected by the cost multiplier or level of pollution. Figure 2 shows the proportion of approved proposals. There are two clear patterns emerging in the data. First, proposals are more likely to be approved under majority than unanimity ($p < 0.001$). Second, proposals are less likely to be approved as the cost multiplier increases ($p < 0.001$). Thus, we reject Hypothesis 5.

Result 5. *The disagreement rate in the Majority treatments is significantly lower than that in the Unanimity treatments. Within each voting rule, the disagreement rate significantly increases by the cost multiplier.*

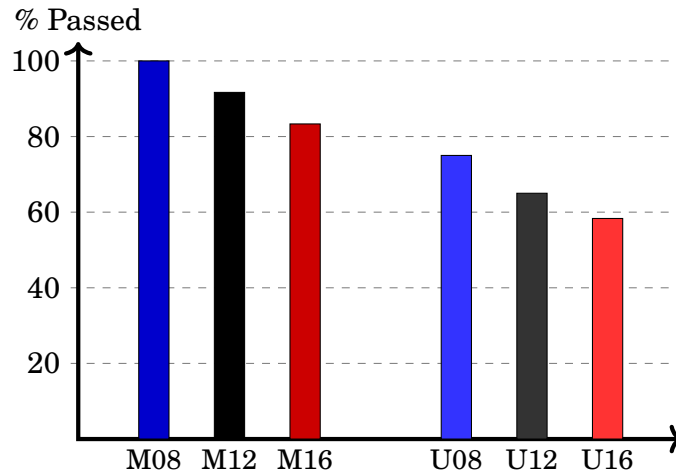


Figure 2: % Proposals Approved by Treatment

7 Conclusions

In this article, we have provided a framework to investigate whether a mechanism in which the burden of a negative externality is shared *ex post* can mitigate or limit socially destructive actions. Our theory highlights that this is possible only under majoritarian agreement rules when the externality costs are high enough. The experimental results show a failure of the unanimous agreement rule in curbing externalities, while only a moderate effect under the majority rule when externality costs are high.

Experiments have traditionally been used in a wide range of settings to understand how key features of the institutions in which decisions are made affect behavior. We do not argue that climate change loss and damage negotiations perfectly resemble our game. However, we believe that the same hurdles that subjects face in the experiment, specifically trying to hold high polluters accountable, are also largely present in the COP27 and COP28 negotiations.

In the experiments, we are able to vary the voting rule in a way that cannot be feasible in actual international negotiations where compliance is voluntary. As discussed, assuming that a majority rule is in place requires acknowledging the existence of an enforcement party which can coerce the non-consenting parties to pay their share. We have shown that even if this was possible, it can still be challenging to hold polluters accountable. The fact that the largest polluters in our experiment often held the proposer role on how to distribute the costs acted as a practical hurdle to implement an equitable sharing of the burden.

Our findings on the overall distribution of costs are unexpected and contrast sharply with the *equitable sharing norm* (Adams, 1963) widely observed in bargaining games

with joint production. In these games, when the surplus to distribute is the result of voluntary contributions or efforts, the sharing of the benefits tends to respect proportionality (Konow, 2000; Gantner et al., 2001; Cappelen et al., 2007), and a similar result holds when the framing of a bargaining game induces different entitlements (Gächter and Riedl, 2005). When distributing endogenous profits, equitable sharing has been shown to foster high levels of efficiency (Baranski, 2016; Dong et al., 2019). However, the same is not true for endogenous costs. More substantively, we show that the effect of fairness norms on efficiency in the gains domain does not necessarily arise in the loss domain. Further research may help illuminate the generality of this result, which, to the best of our knowledge, has not been reported before and has implications beyond climate change loss and damage negotiations.

In providing a tractable model, we aimed to simplify the setting as much as possible by focusing only on the central features that we sought to investigate (voting rule and pollution costs). Naturally, one may wonder what the effect of different negotiation protocols with less structure can be on incentives to impose externalities (Kamm and Siegenthaler, 2022), how the process for selecting the proposer may occur (Lee and Sethi, 2023), as well as the role of player asymmetries in terms of bargaining power (Fréchette et al., 2005b,a; Maaser et al., 2019). These and other questions remain to be answered in future work.

References

- Abatayo, Anna Lou and John Lynham**, “Endogenous vs. exogenous regulations in the commons,” *Journal of Environmental Economics and Management*, 2016, 76, 51–66.
- Adams, J. Stacy**, “Towards an understanding of inequity,” *The Journal of Abnormal and Social Psychology*, 1963, 67 (5), 422.
- Alberti, Federica and César Mantilla**, “A mechanism requesting prices and quantities may increase the provision of heterogeneous public goods,” *Experimental Economics*, 2023.
- Andreoni, James and John Miller**, “Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism,” *Econometrica*, 2002, 70 (2), 737–753.
- Arechar, Antonio A., Simon Gächter, and Lucas Molleman**, “Conducting interactive experiments online,” *Experimental Economics*, 2018, 21, 99–131.

- Baranski, Andrzej**, “Voluntary Contributions and Collective Redistribution,” *American Economic Journal: Microeconomics*, November 2016, 8 (4), 149–173.
- **and Rebecca Morton**, “The determinants of multilateral bargaining: a comprehensive analysis of Baron and Ferejohn majoritarian bargaining experiments,” *Experimental Economics*, 2021.
- Calzolari, Giacomo, Marco Casari, and Riccardo Ghidoni**, “Carbon is forever: A climate change experiment on cooperation,” *Journal of Environmental Economics and Management*, 2018, 92, 169–184.
- Cappelen, Alexander W, Astri Drange Hole, Erik Ø Sørensen, and Bertil Tungodden**, “The pluralism of fairness ideals: An experimental approach,” *American Economic Review*, 2007, 97 (3), 818–827.
- Cherry, Todd L., Peter Frykblom, and Jason F. Shogren**, “Hardnose the dictator,” *American Economic Review*, 2002, 92 (4), 1218–1221.
- Coase, Ronald Harry**, “The problem of social cost,” *The Journal of Law and Economics*, 1960, 3, 1–44.
- Cochard, François, Julie Le Gallo, Nikolaos Georgantzis, and Jean-Christian Tisserand**, “Social preferences across different populations: Meta-analyses on the ultimatum game and dictator game,” *Journal of Behavioral and Experimental Economics*, 2021, 90, 101613.
- Colman, Zack and Karl Mathiesen**, “COP27 Summit Yields ‘Historic Win’ for Climate Reparationsbut Falls Short on Emissions Reductions,” *E&E News*, November 2022.
- Dekel, Sagi, Sven Fischer, and Ro’i Zultan**, “Potential pareto public goods,” *Journal of Public Economics*, 2017, 146, 87–96.
- Dolder, Dennie Van, Martijn J. Van den Assem, Colin F. Camerer, and Richard H. Thaler**, “Standing United or Falling Divided? High Stakes Bargaining in a TV Game Show,” *American Economic Review*, May 2015, 105 (5), 402–07.
- Dong, Lu, Rod Falvey, and Shravan Luckraz**, “Fair share and social efficiency: A mechanism in which peers decide on the payoff division,” *Games and Economic Behavior*, 2019, 115, 209–224.
- Engelmann, Dirk and Martin Strobel**, “Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments,” *American Economic Review*, September 2004, 94 (4), 857–869.

- Essen, Matthew Van and Mark Walker**, “A simple market-like allocation mechanism for public goods,” *Games and Economic Behavior*, 2017, 101, 6–19.
- Fehr, Ernst and Simon Gächter**, “Cooperation and punishment in public goods experiments,” *American Economic Review*, 2000, 90 (4), 980–994.
- Fréchette, Guillaume, John H Kagel, and Massimo Morelli**, “Nominal bargaining power, selection protocol, and discounting in legislative bargaining,” *Journal of Public Economics*, 2005, 89 (8), 1497–1517.
- Fréchette, Guillaume R, John H Kagel, and Massimo Morelli**, “Gamson’s Law versus non-cooperative bargaining theory,” *Games and Economic Behavior*, 2005, 51 (2), 365–390.
- Gächter, Simon and Arno Riedl**, “Moral property rights in bargaining with infeasible claims,” *Management Science*, 2005, 51 (2), 249–263.
- Gantner, Anita, Kristian Horn, and Rudolf Kerschbamer**, “Fair and efficient division through unanimity bargaining when claims are subjective,” *Journal of Economic Psychology*, 2016, 57, 56–73.
- , **Werner Güth, and Manfred Königstein**, “Equitable choices in bargaining games with joint production,” *Journal of Economic Behavior & Organization*, 2001, 46 (2), 209–225.
- Ghidoni, Riccardo, Giacomo Calzolari, and Marco Casari**, “Climate change: Behavioral responses from extreme events and delayed damages,” *Energy Economics*, 2017, 68, 103–115.
- Hamman, John R., Roberto A. Weber, and Jonathan Woon**, “An Experimental Investigation of Electoral Delegation and the Provision of Public Goods,” *American Journal of Political Science*, 2011, 55 (4), 738–752.
- Hauser, Oliver P, David G Rand, Alexander Peysakhovich, and Martin A Nowak**, “Cooperating with the future,” *Nature*, 2014, 511 (7508), 220–223.
- Kamm, Aaron and Simon Siegenthaler**, “Commitment timing in coalitional bargaining,” *Experimental Economics*, 2022, pp. 1–25.
- Kim, Duk Gyoo and Wooyoung Lim**, “Multilateral Bargaining over the Division of Losses,” CESifo Working Paper Series 8011, CESifo Group Munich Dec 2019.

- Konow, James**, “Fair shares: Accountability and cognitive dissonance in allocation decisions,” *American Economic Review*, 2000, 90 (4), 1072–1091.
- Lee, Natalie and Ravideep Sethi**, “Recognition probability in legislative bargaining,” *Journal of Experimental Political Science*, 2023, pp. 1–22.
- Maaser, Nicola, Fabian Paetzel, and Stefan Traub**, “Power illusion in coalitional bargaining: An experimental analysis,” *Games and Economic Behavior*, 2019, 117, 433–450.
- Mora, Camilo, Daniele Spirandelli, Erik C. Franklin, John Lynham, Michael B. Kantar, Wendy Miles, Charlotte Z. Smith, Kelle Freel, Jade Moy, Leo V. Louis et al.**, “Broad threat to humanity from cumulative climate hazards intensified by greenhouse gas emissions,” *Nature climate change*, 2018, 8 (12), 1062–1071.
- Nordhaus, William**, “Climate clubs: Overcoming free-riding in international climate policy,” *American Economic Review*, 2015, 105 (4), 1339–1370.
- Ostrom, Elinor**, “The value-added of laboratory experiments for the study of institutions and common-pool resources,” *Journal of Economic Behavior & Organization*, 2006, 61 (2), 149–163.
- , **Roy Gardner, and James Walker**, *Rules, games, and common-pool resources*, University of Michigan press, 1994.
- Pevnitskaya, Svetlana and Dmitry Ryvkin**, “The effect of access to clean technology on pollution reduction: An experiment,” *Games and Economic Behavior*, 2022, 136, 117–141.
- Pigou, Arthur C.**, *The Economics of Welfare*, 4th ed., London: Macmillan, 1932.
- Rodriguez-Lara, Ismael and Luis Moreno-Garrido**, “Self-interest and fairness: self-serving choices of justice principles,” *Experimental Economics*, 2012, 15 (1), 158–175.
- Rodriguez-Sickert, Carlos, Ricardo Andrés Guzmán, and Juan Camilo Cárdenas**, “Institutions influence preferences: Evidence from a common pool resource experiment,” *Journal of Economic Behavior & Organization*, 2008, 67 (1), 215–227.
- Sengupta, Somini**, “Calls for Climate Reparations Reach Boiling Point in Glasgow Talks,” *The New York Times*, November 2021.
- Stoddard, Brock, James M. Walker, and Arlington Williams**, “Allocating a voluntarily provided common-property resource: An experimental examination,” *Journal of Economic Behavior & Organization*, 2014, 101, 141–155.

Stranlund, John K., James J. Murphy, and John M. Spraggon, “An experimental analysis of compliance in dynamic emissions markets,” *Journal of Environmental Economics and Management*, 2011, 62 (3), 414–429.

A Proofs

Proof of Lemma 1: If $p_i \leq \frac{1}{2}$, then the payoff of accepting the proposal is $g_i - p_i C$, which yields greater payoffs than $g_i - \frac{C}{2}$. If player i votes for the proposal, then p_i has to be less than $\frac{1}{2}$. When voting for the proposal with $p_i > \frac{1}{2}$, the player's voting behavior yields either a strictly lower payoff than rejecting the proposal. \square

Proof of Lemma 2: Since all other players j would accept any proposal $p_j \leq \frac{1}{2}$ (Lemma 1), the proposer maximizes her payoff by choosing $p_j = \frac{1}{2}$ for all $j \neq i$. Depending on the voting rule q , other forms of the proposal are also possible, but it must not involve a positive amount of cost to the proposer because such proposal is strictly dominated. \square

Proof of Proposition 1: In the second stage, non-proposer j votes for the proposal by Lemma 1. Proposer i does not want to propose differently because offering less costs to others will lower her payoffs (Lemma 2), and offering more costs will lead to the rejection of the proposal, which will again lower her payoffs. In the first stage, given that two other players claim g_j^* and g_k^* respectively, player i 's payoff when claiming g_i is $g_i - 0$ with probability $\frac{1}{3}$, and $g_i - \frac{\alpha(g_i + g_j^* + g_k^*)}{2}$ with probability $\frac{2}{3}$. The expected payoff is then $\frac{1}{3}g_i + \frac{2}{3}(g_i - \frac{\alpha(g_i + g_j^* + g_k^*)}{2}) = (1 - \frac{\alpha}{3})g_i - \alpha \frac{g_j^* + g_k^*}{3}$, which increases monotonically in g_i because $\alpha < 3$. Thus, the dominant strategy in the first stage is to choose $g_i^* = E$. Since the proposer and the two non-proposers vote for the proposal, the proposal is approved. This equilibrium is unique up to a permutation of the players' identities. \square

Proof of Proposition 2: Let $p_j > p_k$ where j and k are voters. Then, player j votes against and player k in favor, because she receives a cost that is lower than her outside option. If $p_j = p_k = C/2$, then both player vote in favor. As such, the proposal always passes. Clearly, the proposer assigns herself $p_i = 0$ because there is paying for any costs can only decrease her payoff and leave unchanged the probability of a proposal being approved. \square

Proof of Proposition 3: In the second stage, the proposer and the non-proposer who take zero costs vote for the proposal because $E - 0 > E - \alpha \frac{3E}{2}$. Although the non-proposer who are burdened with the entire costs of $\alpha 3E$ votes against it, the proposal is approved since two players vote for it. In the first stage, given that two other players claim g_j^* and g_k^* respectively, player i 's payoff when claiming g_i is $g_i - 0$ with probability $\frac{2}{3}$, and $g_i - \alpha(g_i + g_j^* + g_k^*)$ with probability $\frac{1}{3}$. The expected payoff is then $\frac{2}{3}g_i + \frac{1}{3}(g_i - \alpha(g_i + g_j^* + g_k^*)) = (1 - \frac{\alpha}{3})g_i - \alpha \frac{2(g_j^* + g_k^*)}{3}$, which increases monotonically in g_i because $\alpha < 3$. Thus, the dominant strategy in the first stage is to choose $g_i^* = E$. \square

Proof of Proposition 4: In the second stage, there are zero costs to share given $g_i^* = 0$ for all i . Since every player is allocated zero costs, the proposal is trivially approved. The

payoffs on the equilibrium path are zero to each player. Given two other players choose $g_j^* = 0$ in the first stage, player i 's payoff when claiming $g_i > 0$ is g_i with probability $\frac{1}{3}$ because player i takes zero costs to herself (Lemma 1), and $g_i - \alpha g_i$ with probability $\frac{2}{3}$ because the other two players assign the entire costs to player i when one of them becomes the proposer. The expected payoff is then $(1 - \frac{2}{3}\alpha)g_i$, which decreases monotonically in g_i when $\alpha > \frac{3}{2}$. Thus, the best response to $g_j^* = 0$ for $j \neq i$ is to choose $g_i^* = 0$. \square

B Robustness of Results

In this appendix, we show the results of linear regression analyses of what we reported in the main text.

Table B.1: The determinants of claims (pollution)

Claim	All		Majority		Unanimity	
	(1)	(2)	(3)	(4)	(5)	(6)
CostMultiplier	-54.93*** (8.73)	-51.90*** (10.04)	-53.67*** (11.34)	-52.62*** (13.11)	-56.19*** (13.34)	-52.09*** (16.16)
Period	5.94*** (1.11)	5.50*** (1.18)	2.99* (1.36)	2.51 (1.48)	8.88*** (1.70)	8.39*** (1.79)
Majority	0.93 (5.67)	-1.69 (6.65)				
RiskAversion		-3.91 (2.50)		-4.17 (3.45)		-3.69 (3.73)
Female		6.39 (7.03)		4.91 (8.46)		7.55 (11.21)
Age		2.04* (1.01)		0.48 (1.41)		3.28 (1.54)
#Friends		1.45 (5.17)		8.58 (9.40)		1.13 (7.29)
N	960	875	480	430	480	445
R ²	0.118	0.122	0.122	0.121	0.125	0.142

OLS regression of claim. Majority is the indicator of the Majority treatments. RiskAversion is a measure of risk aversion based on the two answers from the post-experiment survey, varying from 1 (most averse) to 4 (least averse). Observations from whom preferred not to answer their gender and age are omitted. #Friends is an indicator whether there are friends in the same session. The standard errors clustered at the individual level are in parentheses. *, **, and *** indicate statistical significance at the 5% level, 1% level, and 0.1% level, respectively.

Table B.2: The treatment effect of voting rule

Claim	$\alpha = 0.8$		$\alpha = 1.2$		$\alpha = 1.6$	
	(1)	(2)	(3)	(4)	(5)	(6)
Majority	-4.44 (3.65)	-7.20 (4.64)	6.45 (8.09)	5.04 (9.04)	-1.02 (12.56)	-4.91 (12.89)
Period		2.30 (1.21)		6.26** (2.29)		6.75** (1.98)
RiskAversion		-1.00 (2.14)		-0.65 (3.20)		-7.74 (5.05)
Female		3.48 (5.28)		-0.59 (10.10)		15.87 (13.71)
Age		1.03 (0.48)		2.70 (1.62)		2.47 (2.16)
#Friends		3.14 (3.57)		-2.14 (8.01)		8.38 (13.45)
N	240	210	360	325	360	445
R^2	0.010	0.059	0.004	0.053	0.000	0.050

OLS regression of claim by cost multiplier. The standard errors clustered at the individual level are in parentheses. *, **, and *** indicate statistical significance at the 5% level, 1% level, and 0.1% level, respectively.

Table B.3: The determinants of approval

Voting Result	(1)	(2)	(3)
Majority	0.26*** (0.05)		0.26*** (0.05)
CostMultiplier		-0.21*** (0.08)	-0.21** (0.08)
Period			0.03 (0.02)
N	320	320	320
R^2	0.095	0.024	0.127

OLS regression of voting outcome. Since each group of three has one voting outcome, the group-level observations are used, and the individual characteristics are not controlled. The standard errors clustered at the individual level are in parentheses. *, **, and *** indicate statistical significance at the 5% level, 1% level, and 0.1% level, respectively.

Table B.4: The determinants of Proportion of Costs Offered, $\alpha = 1.6$

	M16		U16	
	(1)	(2)	(3)	(4)
Pollution (relative)	0.31** (0.09)	0.54*** (0.11)	0.14* (0.05)	0.17 (0.09)
Share to self (0 or 1)	-0.16*** (0.03)	0.07 (0.07)	-0.09*** (0.02)	-0.07 (0.05)
Pollution \times Share to self		-0.69*** (0.18)		-0.08 (0.16)
N	540	540	540	540
R^2	0.069	0.090	0.076	0.078

OLS regression of offered share of the costs. The unit of observation is a share of the costs offered by a subject to each member of the group. *Pollution* is the relative size of the claim in the group for a given recipient. *Share to self* is the indicator whether the offered share is to herself. The standard errors clustered at the individual level of the subject making the offer are in parentheses. *, **, and *** indicate statistical significance at the 5% level, 1% level, and 0.1% level, respectively.

Table B.5: The determinants of voter behavior

Yes Vote	Linear		Logit	
	(1)	(2)	(3)	(4)
OwnShare	-0.92*** (0.05)	-0.91*** (0.05)	-6.13*** (0.87)	-6.10*** (0.96)
ProposerShare	0.51*** (0.14)	0.55** (0.16)	2.78** (0.99)	3.08** (1.12)
Majority	-0.07 (0.04)	-0.08 (0.05)	-0.48 (0.25)	-0.53 (0.28)
CostMultiplier	-0.12* (0.05)	-0.08 (0.05)	-0.70 (0.38)	-0.43 (0.40)
Period		0.01 (0.01)		0.08 (0.08)
RiskAversion		-0.01 (0.02)		-0.06 (0.11)
Female		0.02 (0.04)		0.15 (0.28)
Age		0.00 (0.00)		0.03 (0.04)
#Friends		0.01 (0.03)		0.10 (0.23)
N	640	581	640	581
(Pseudo-) R^2	0.361	0.364	0.361	0.352

OLS regression of voter behavior. Observations from non-proposers are used. OwnShare is the proposed share of the costs, and ProposerShare is the share of the costs proposed to the proposer herself. The standard errors clustered at the individual level are in parentheses. *, **, and *** indicate statistical significance at the 5% level, 1% level, and 0.1% level, respectively.

C Experimental Instructions

(Instructions for Majority, Cost multiplier 1.2)

Welcome to this experiment. During the experiment, please do not close this window or leave the web pages in any other way. If you do close your browser or leave the task, you will not be able to re-enter, and we will not be able to pay you. It is therefore important that you complete this experiment without interruptions. If you have questions regarding the procedure of the experiment or want to troubleshoot, please contact the experimenter.

Please read the instructions carefully. There will be a quiz to check your understanding of the instructions. The cash payment you will receive at the end of the experiment will depend on the decisions you make as well as the decisions other participants make. The currency in this experiment is called "tokens."

Overview

In this experiment, you will be placed in a group of three people and will engage in three main stages. First, each member of the group will have a chance to claim between 0 and 200 tokens for him or herself. For each token you take, you generate a cost for the group. In stage two, each person will propose a division of the group's total cost. In the third stage, proposals are voted up or down by majority rule.

Stage 1. Claim tokens

Everyone is initially given 1,000 tokens in a private account. There is also a public account with 600 tokens, and you can claim up to 200. The claimed tokens are added to your private account. For each token that you claim, you generate a cost of 1.2 tokens to the group. You will deal with the total costs incurred in your group in Stage 2.

Stage 2. Make a proposal

You will observe who claimed how much, and accordingly, the total costs. Each member of the group will propose a division of the costs to be paid, by typing the costs allocated to each member.

	Member 1	Member 2	Member 3
Costs allocated (in Tokens)	_____	_____	_____

The sum of the allocated costs must be equal to the total costs. After all group members submit their proposal, one of the three proposals will be randomly selected with equal probability to be voted on.

Stage 3. Vote Up or Down

Examine the chosen proposal. Vote up or down the proposal. Each member has one vote.

- The proposal is approved when two or more members vote for it. The tokens allocated to you are DEDUCTED from your private account. In this case, your payoff in this period is:

$$1000 + [\#Tokens\ you\ claimed] - [\#Tokens\ allocated\ to\ you\ in\ the\ approved\ proposal]$$

- Otherwise, that is, if two or more members votes against the proposal, HALF OF THE TOTAL TOKENS will be DEDUCTED from every member's private account. In this case, your payoff in this period is:

$$1000 + [\#Tokens\ you\ claimed] - [Half\ of\ the\ total\ costs]$$

Example

(*This is only for illustration. Any numbers used in this example do not intend any guidance.)

For illustration, suppose Members 1, 2, and 3 claim 0 tokens, 100 tokens, and 200 tokens, respectively. Then, Members 1, 2, and 3 generate a cost of 0 tokens, 120 tokens, and 240 tokens each.

In Stage 2, each member proposes a division of the total costs, 360 tokens.

In Stage 3, the randomly selected proposal is put up to the vote. If everyone votes for the proposal, the costs are allocated as proposed.

If at least one member votes against the proposal, half of the costs (180 in this example) are allocated to each member.

How the Groups are Formed You will participate in a total of 10 periods consisting of Stages 1–3 described above. In each period, all participants will be randomly assigned to new groups of three members. Each member of a group will have an ID number (from 1 to 3). Since IDs will be reassigned as well, everyone remains anonymous, ever after the end of the experiment.

Information Feedback

You will be provided with a summary of what happened in the period, including the selected proposal for distributing the costs, the proposer's ID, the voting outcome, and your payoff from the period.

Payment

The server computer will randomly select one of the 5 periods you have participated in, and your payoff in that period will be paid. Each period has an equal chance to be chosen for the final cash payment, so it is in your interest to take each period equally seriously. Your payoff in the selected period is converted to KRW at the rate of 1 Token

= 14 KRW.

Summary of the process

1. The experiment consists of 5 periods.
2. In each period, every participant has a private account with 1,000 tokens and will be randomly grouped with two other participants. Each member of the group assigns and ID number.
3. In Stage 1 of each period, you decide how many tokens (up to 200) to claim from a group account. Every token you claim will generate a cost of 1.2 tokens to the group. The tokens you claim accrue to your private account.
4. In Stage 2 of each period, each member will observe how many tokens other members claimed and the total costs of the group. Each member then submits a proposal to divide the costs to each member.
5. In stage 3 of each period, one of the three submitted proposals is randomly selected, and you vote for or against the chosen proposal. If the proposal is approved with two or more yes votes, the tokens allocated to you are deducted from your account. If the proposal is rejected, half of the total costs are deducted from your account.

Comprehension check The following questions are provided to check your understanding of the instructions. If you want to read the instructions again, please click on [Back to Instructions]. You will move on to the next page only after you answer all questions correctly.

- Q1. Imagine the following situation. You claimed 200 tokens, and the other two group members claimed 100 tokens each. What are the total costs to be paid?
- Q2. Suppose three group members claimed 0 tokens, 100 tokens, and 200 tokens. Recall that everyone starts with an endowment of 1,000 tokens. Which of the following is incorrect?
1. Depending on the situation, one could earn a payoff less than 1,000 tokens.
 2. In this situation, the total costs are 360 tokens.
 3. For this decision round, the payoff of the member who claimed 0 tokens is 1,000 tokens in any case.
 4. The chosen proposal is approved only when all the three group members vote for it.

Q3. Which of the following is correct?

1. Whatever happened in the previous periods will not affect the formation of the new groups and the selection of the proposer.
2. Although your group members are anonymous, they are the same for the entire 10 periods.
3. The sum of the payoffs in the entire 10 periods will be paid.
4. Your decisions in the previous periods can make the following periods favorable to you.