Multilateral Bargaining over the Division of Losses*

Duk Gyoo Kim† Wooyoung Lim‡

June 22, 2021

Abstract

Many-player divide-the-dollar games have been a workhorse in the theoretical and experimental analysis of multilateral bargaining. If we are dealing with a loss, that is, if we consider many-player "divide-the-penalty" games for, e.g., the location choice of noxious facilities, the allocation of burdensome chores, or the reduction of carbon dioxide emissions at a climate change summit, the theoretical predictions are not simply those from the divide-the-dollar games with the sign flipped. We show that the stationary subgame-perfect equilibrium (SSPE) is no longer unique in payoffs. The most "egalitarian" equilibrium among the stationary equilibria is a mirror image of the essentially unique SSPE in the Baron-Ferejohn model. That equilibrium is fragile in the sense that allocations are sensitive when responding to changes in parameters, while the most "unequal" equilibrium is not affected by such changes. Experimental evidence clearly supports the most unequal equilibrium: Most of the approved proposals under a majority rule involve an extreme allocation of the loss to a few members. Other observations such as no delay, proposer advantage, and the acceptance rate are also consistent with the predictions based on the most unequal equilibrium.

JEL Classification: C78, D72, C92

Keywords: Multilateral bargaining, Loss division, Laboratory experiments

*We thank Andrezj Baranski, Peter Duersch, Hülya Eraslan, Hans-Peter Grüner, Bård Harstad, Pablo Hernandez-Lagos, Kai Konrad, Marco Lambrecht, Joosung Lee, Robin S. Lee, Shih En Lu, Rebecca Morton, Nikos Nikiforakis, Salvatore Nunnari, Satoru Takahashi, Huan Xie, and the conference and seminar participants at HeiKaMaxY workshop, NYU Abu Dhabi, Mannheim, Max Planck Institute for Tax Law and Public Finance, UBC, Victoria, UMass Amherst, Queen's, Carleton, Concordia, McMaster, Florida State, Toronto, Western Ontario, Guelph, Bilkent, Yonsei, 2019 International Meeting on Experimental and Behavioral Social Sciences, 2019 North American Meeting of the Econometric Society, 2019 Economic Science Association World Meeting, 2019 Korean Economic Review International conference, 2019 Canadian Public Economics Group Conference, 2019 Southern Economic Association Conference, and 2020 Econometric Society World Congress for their comments and suggestions. This study is supported by a grant from the Research Grants Council of Hong Kong (Grant No. GRF-16500318).

†Department of Economics, University of Mannheim, *d.kim@uni-mannheim.de*

‡Department of Economics, The Hong Kong University of Science and Technology, *wooyoung@ust.hk*

1 Introduction

Multilateral bargaining refers to a situation in which a group of agents with conflicting interests try to bargain under a predetermined voting rule. Many-player divide-the-dollar (henceforth, DD) games where a group of agents reach an agreement on a proposal dividing a dollar have served well as an analytic tool for understanding multilateral bargaining behavior (Baron and Ferejohn, 1989). However, we claim that this model sheds light on only one side of multilateral bargaining: The other side addresses the distribution of a loss or a penalty. Our contribution is twofold: We (1) demonstrate that multilateral bargaining over the distribution of bads is theoretically different from that over goods and (2) provide experimental evidence that clearly diverges from the standard findings in the experimental multilateral bargaining literature.

Real-life situations dealing with the distribution of a loss are as common as those addressing a surplus. Taxation for public spending could be understood as a distribution of burdens. A location choice of a noxious facility is an example of the allocation of a loss, as those closer will suffer more from the disutility of the facility than those in other areas. The climate change summit is another example of dividing a penalty in the sense that the participating countries share the global consensus on the need to reduce carbon dioxide emission levels, but no single country wants responsibility for the whole burden, as it may be harmful to their economic growth. Despite its relevance to many policy issues, little attention has been paid to multilateral bargaining over the division of losses.[1] Such inattention might be due to a conjecture that the theoretical predictions of a many-player "divide-the-penalty" (henceforth, DP) game would be exactly the inverse of those of the DD game. We claim that this is not the case, and our claim does not rely on any behavioral/psychological assumptions, including loss aversion.

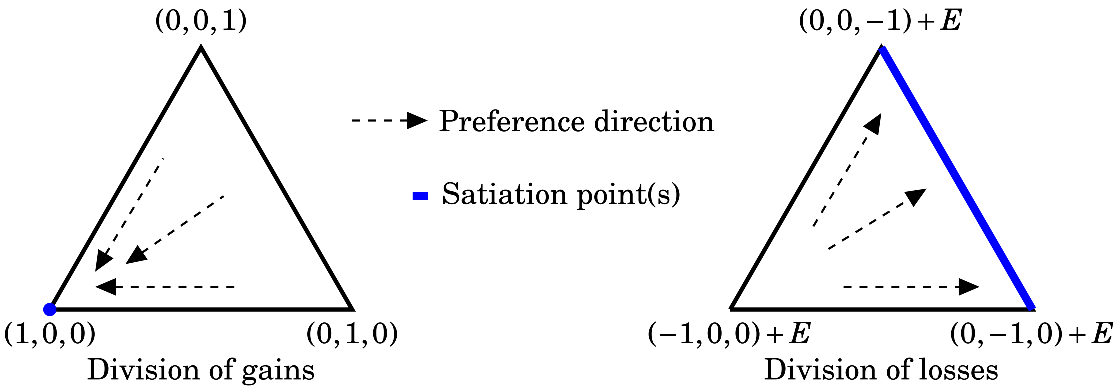


Figure 1: Different Preference Directions and Satiation Points

Figuratively speaking, comparing the DD game with the DP game is *not* analogous to comparing the allocation of a "half-full" cup of water to that of a "half-empty" cup; instead, it is analogous to

[1] Some studies, including Aghion and Bolton (2003), Aghion et al. (2004), Harstad (2005), and Harstad (2010), examine coalition formation games concerning the participation constraint. In that the core idea of those studies is related to the question of how the cost of a public project should be split among participants, the division of losses itself is not a new question. Multilateral bargaining games consider both coalition formation and the allocation of resources to the coalition.

comparing the allocation of a full cup of "clean" water when everyone is thirsty to that of a full cup of "filthy" water when everyone is fully saturated. The latter example deals with fundamentally different objectives that have the opposite preference directions. A 2-dimensional unit simplex, which captures the allocation of resources (normalized to one) among three players, can also illustrate this analogy. In Figure 1, player 1 at the bottom-left vertex has a unique satiation point over the division of gains, but she prefers any linear combination of the other two vertices over the division of losses. Although the procedure for dividing a fixed amount of resources would be identical in both situations, the preference directions on the object are not merely flipped. This difference is not due to the domain of utilities: Even if every subject is endowed with an external payoff E that is sufficient to enjoy a positive level of utility overall, the division of losses is still different from the division of gains regardless of the size of E.[2] Thus, the fundamental difference does not rely on the framing effect on the loss.

Another key difference comes from a proposer advantage in the division of losses: Whoever holds the position with stronger bargaining power cannot seize an advantage that is greater than gaining zero losses. In the DD game, a proposer exploits rent from being the proposer by forming a minimum winning coalition (MWC) to the extent that the number of "yes" votes is just sufficient for the proposal to be approved and by offering the members in the MWC their continuation value so that rejecting the offer would not make them better off. Altogether, a significant amount of the proposer advantage is predicted in the DD game. However, the proposer in the DP game, who will at best enjoy no losses, may not be better off than those in the MWC, who could also enjoy no losses. In other words, one could imagine that a proposer in the DD game says, "I want to have as much gain as possible, so there is no way to give my coalition members more than their minimum willingness to accept." In contrast, a proposer in the DP game says, "I want to have as little loss as possible, but there are many ways to assign losses to my coalition members less than their maximum willingness to suffer."

The fact that the proposer cannot enjoy an advantage greater than zero losses is a source of the primary theoretical difference between the DD game and the DP game. While the DD game has a unique stationary subgame-perfect equilibrium (SSPE) in payoffs (Eraslan, 2002),[3] the DP game has a continuum of stationary subgame-perfect equilibria with different ex-post distributions of the payoffs. The strategy of one SSPE, which we call the utmost inequality (UI) equilibrium, is for the proposer to assign the total penalty to one randomly chosen member: The other members without a penalty will accept the proposal because the continuation value (the expected payoff from moving on to the next bargaining round) would be strictly smaller than 0. At the other extreme, the strategy of another SSPE, which we call the most egalitarian (ME) equilibrium, is for the proposer to distribute the penalty across all of the members except herself such that MWC members will be indifferent between

[2] Our experimental design is based on the implementation of this idea.

[3] More specifically, with the identical players and the same recognition probability, the SSPE on the DD game has a unique ex-post distribution of payoffs. For example, with three players, no discount, and a simple majority rule, the distribution of payoffs is (2/3, 1/3, 0), although we do not know the identities of the proposer and the MWC member. Eraslan and McLennan (2013) extend Eraslan (2002) by showing the uniqueness of the expected payoffs of the stationary equilibria in more general games. In the DP game considered in this paper, the ex-ante expected payoff is simply $-\delta/n$ in any of the stationary subgame-perfect equilibria, so it is still unique.

accepting and rejecting the current offer. Of course, any intermediate strategy between these two extreme SSP equilibrium strategies can constitute an SSPE. Therefore, the primary goal of this paper is to comprehensively investigate the DP game and compare it with the DD game both theoretically and experimentally.

Laboratory experiments have been a useful tool in the multilateral bargaining literature. We claim that the use of lab experiments is more critical for the DP game. Even if we narrow down our focus to stationary strategies, theory is silent in guiding us toward the equilibrium that is more likely to be consistent with our observations. Anecdotal empirical evidence might be sporadically available, but we cannot be free from the issues of measurement, endogeneity, and unobservable heterogeneities to identify a clear causal link. Moreover, it is challenging, if not impossible, for experimenting policymakers to test different situations where an actual loss would be distributed.

Among the many potential directions for experiment design to test the theory, we chose the simplest possible, yet most frequently revisited, ones. We conducted experiments on four treatments that vary by two dimensions: the group size (either 3 or 5) and the voting rule (either majority or unanimity). Theoretical predictions based on the ME equilibrium were used as null hypotheses because this equilibrium resembles the essentially unique SSPE in the DD game, and it approaches the unique equilibrium under unanimity as the qualified number of voters for approval goes to the group size. Experimental evidence clearly rejects the ME equilibrium. Instead, the UI equilibrium is the most consistent with our experimental observations. Most of the approved proposals under a majority rule involve an extreme allocation of the loss to a few members. That is, in three-member bargaining, one member exclusively receives all the losses, and in five-member bargaining, either one member receives the total loss or two members receive a half each. The utilitarian efficiency, meaning no delay in reaching an agreement, and the proposer advantage are well observed.

We claim that our experimental evidence is a watershed that identifies potential to extend this study. The experimental results of the DD game are overall in line with the theoretical predictions.[4] If the observed patterns were close to the mirror image of the SSPE over the division of gains, although we identify other SSP equilibria over the division of losses, we could conclude that the many-player divide-the-dollar game is still sufficiently appropriate for studying multilateral bargaining over the division of losses. Since we find crucial differences both theoretically and experimentally, it is worth revisiting all the important studies on multilateral bargaining where the main motivating situations focus on the division of losses.

The rest of this paper is organized as follows. In the following subsection, we discuss the related literature. Section 2 presents the model of the divide-the-penalty game, and Section 3 describes the

[4]According to Palfrey (2016), three major predictions of the essentially unique SSPE in the DD game are (1) the formation of the minimum winning coalition, (2) utilitarian efficiency (that is, no efficiency loss due to delays), and (3) proposers' full rent extraction. All of those predictions are consistent with the experimental findings. Perhaps the proposer advantage is not remarkably consistent with the theoretical predictions as proposers tend to take a smaller advantage than they can fully exploit (Fréchette et al., 2003). Still, several follow-up studies, including Agranov and Tergiman (2014) and Baranski and Kagel (2015) provide evidence suggesting that the proposers' partial rent extraction is mainly due to uncertainties particularly existing in the laboratory setting, which could be mitigated via anonymous pre-bargaining communications.

theoretical properties of the model. The experimental design, hypotheses, and procedure are discussed in Section 4. We report our experimental findings in Section 5. Section 6 discusses further issues, and Section 7 concludes the paper.

1.1 Related Literature

This study stems from a large body of literature on multilateral bargaining. A legislative bargaining model initiated by Baron and Ferejohn (1989) has been extended (Eraslan, 2002; Norman, 2002; Jackson and Moselle, 2002), adopted for use with more general models (Battaglini and Coate, 2007; Diermeier and Merlo, 2000; Volden and Wiseman, 2007; Bernheim et al., 2006; Diermeier and Fong, 2011; Ali et al., 2019; Kim, 2019), and experimentally tested (Diermeier and Morton, 2005; Fréchette et al., 2003, 2005; Fréchette et al., 2012; Agranov and Tergiman, 2014; Kim, 2018).[5] Our contribution to this literature is to show that the theoretical predictions of the DP game under a majority rule could be significantly different due to the natural restriction of the proposer advantage: In the DP game, the maximum advantage available to the proposer is to receive no penalties.[6]

In that the fundamental idea of the model relates to the allocation of bads, this study is pertinent to chore division models (Peterson and Su, 2002), a subset of envy-free fair division problems (Stromquist, 1980) in which the divided resource is undesirable. Social choice theorists are well aware of the distinctive difference between the allocation of goods and that of bads. Bogomolnaia et al. (2018) show that in the division of bads, unlike that of goods, no allocation rule dominates the others in a normative sense. While the literature on envy-free division has focused more on the algorithms or protocols that lead to the desired allocation, this paper only considers predetermined voting rules and does not focus on the design of algorithms. Another area of the literature philosophically connected to our study is those works addressing the principle of equal sacrifice in income taxation (Young, 1988; Ok, 1995) in which the primary purpose is to justify the traditional equal sacrifice principles in taxation from a non-utilitarian perspective by showing that the utility function satisfying equal sacrifice principles could be a consequence of more primitive concepts of distributive justice. Although taxation for public spending or redistribution is related to the idea of distributing monetary burdens, we try not to be normative in this paper. Experimental findings on multilateral bargaining over the division of losses are rare. Gaertner et al. (2019) find that when subjects are endowed with a different amount of money and collectively determine the allocation of a loss under a unanimity rule, the proportionality principle—resource allocation proportional to the endowment—is rarely observed.

This study is also remotely connected to the literature documenting behavioral asymmetries between the gain and loss domains. From the many studies about loss aversion, we know that human behavior when dealing with losses is different from that when experiencing gains. In this regard,

[5] For more complete review, see Eraslan and Evdokimov (2019).

[6] In the sense that only one voting rule—unanimity—is considered in two-person bargaining, our analysis under a unanimity rule is related to Weg and Zwick (1991). They provide experimental evidence that there are no differences between the division of gains and losses, which is consistent with our findings under a unanimity rule treatment.

Christiansen and Kagel (2019) is one study philosophically related to ours: They examine how the framing changes three-player bargaining behavior. In particular, based on the model studied by Jackson and Moselle (2002), they study two treatments that are isomorphically the same in theory but framed differently. Since the theoretical predictions of the two treatments are identical, their primary purpose is to observe the framing effect.[7] Their study is rather related to the literature on the discrepancies between willingness-to-pay and willingness-to-accept. The crucial difference between our study and theirs is that we deal with different incentive structures, so the framing does not play an important role. The experimental design considered in Christiansen and Kagel (2019) can be regarded as a comparison of 'half-full' versus 'half-empty' glass of water, figuratively speaking. In the sense that we indirectly compare an economic outcome on a gain domain with that on a loss domain, Gerardi et al. (2016) is another closely related study. They compare the penalty of not turning out to vote with a lottery for those who do turn out, show that these two incentive structures are theoretically similar, and provide experimental evidence that voters are more likely to turn out under a lottery treatment than under a penalty treatment.

Although it may appear that the comparision between public bad prevention and public good provision (Andreoni, 1995) is somewhat related, the comparison between the DD game and the DP game is distinctively different from the comparison between public good provision and public bad prevention because the former does not involve any form of externality. Regarding the treatment of public bads, a political economy examination of the NIMBY ("not in my back yard") conflict could also be related to this paper. Levinson (1999) demonstrates that local taxes for hazardous waste disposal can be inefficient because of the tax elasticity of polluters' responses. Fredriksson (2000) shows that a centralized system for siting hazardous waste treatment facilities is suboptimal compared to the decentralized system because of lobbying activities. Feinerman et al. (2004) adopt a model of a competitive real estate market between two cities and provide suggestive evidence that if all cities in the region form political lobbies, the political siting is geographically close to the socially optimal location. To the best of our knowledge, the political procedure and the equilibrium outcomes under a qualified voting rule have not been investigated in previous studies.

2 A Model

We consider a many-player divide-the-penalty game. As the many-player divide-the-dollar game à la Baron and Ferejohn (1989) aims to understand multilateral bargaining over a surplus, the divide-the-penalty game will serve as a theoretical tool to understand multilateral bargaining over a loss.

There are n (an odd number greater than or equal to 3) players indexed by $i \in N = \{1, \ldots, n\}$. A feasible allocation share is $p = (p_1, \ldots, p_n) \in \{[-1,0]^n | \sum_i p_i = -1\}$, and the set of feasible allocation shares is denoted as P. We consider a q-quota voting rule: The consent of at least $q \leq n$ players is

[7] Christiansen et al. (2018) continue examining the framing effect using the Baron-Ferejohn model, in addition to the role of communication, as in Agranov and Tergiman (2014) and Baranski and Kagel (2015).

required for a proposal to be approved. The voting rule is called a dictatorship if $q = 1$, a (simple) majority if $q = \frac{n+1}{2}$, unanimity if $q = n$, and a super-majority if $q \in \{\frac{n+3}{2}, \dots, n-1\}$.

The amount of the loss increases as time passes, so delay is costly. The cost of delay is captured by the growth rate of the loss, $g \in [1, \infty)$ per delay. At the same time, delay also dilutes the disutility of the penalty. If players prefer having the disutility tomorrow to having the same amount of disutility today, they may want to postpone the actual allocation of the penalty as long as possible, so that the disutility of the allocation can be diluted. Let $\beta \in (0,1]$ denote such time preference. When the allocation of the penalty is made in round t, player i's utility is $U_i^t(p) = (\beta g)^{t-1} p_i$. For notational convenience, let $\delta \equiv \beta g$, which can be larger or smaller than 1.[8] Over the division of losses, these two factors, β and g, lead to different incentives. When the time preference dominates the growth rate of the penalty, that is, when $\delta < 1$, players have an incentive to postpone the actual allocation of the loss. Otherwise, players want to make a decision as quickly as possible. We focus on $\delta \geq 1$ because it can capture more pertinent situations: If the nature of bargaining drives the relevant parties to postpone their agreement as long as possible, such bargaining may deal with relatively trivial issues.[9] To complete the model, we assume that each player earns the utility of -1, the lowest possible static payoff, when they do not reach an agreement for infinite rounds of bargaining. This assumption merely corresponds to the assumption in the DD game in which each player earns 0, the lowest possible static payoff, when disagreeing forever.[10] It is worth noting that the utility of the infinite disagreements does not need to be negative one. It is sufficient for it to be smaller than the ex-ante expected payoff in the second round, $-\delta/n$, and the equilibria are sustained even when $\delta < 1$ so that the loss exponentially decreases to zero.[11] In this regard, the practical upper bound of loss in the experiment (due to the minimum guaranteed payment), as well as alternative assumptions about infinite disagreements to match the real-world examples, do not affect the characterization of the equilibria.

Players bargain over the loss until they reach an agreement. The timing of the game is as follows:

1. In round $t \in \mathbb{N}_+$, one player is randomly recognized as a proposer with equal probability. The selected player proposes an allocation of $-g^{t-1}$ in terms of proportions.

2. Each player votes on the proposal. If it is approved, that is, if q or more players accept the proposal, the proposal is implemented, $U_i^t(p)$ is accrued, and the game ends. If the proposal is not accepted, the game moves on to round $t+1$.

3. In round $t+1$, a player is randomly recognized as the proposer. The game repeats at $t+1$.

[8] A discount factor in the standard dynamic models, $\delta \in [0,1)$, can be understood as the depreciation rate (the inverse of the growth rate), $1/g$, times the subjective time-discount factor, β. In this case, δ is always smaller than 1, so the distinction between the depreciation rate and time preference is not crucial.

[9] The stationary subgame-perfect equilibria are sustained in $\delta < 1$, but the equilibria involve socially inefficient outcomes. More details with $\delta < 1$ are discussed in Section 6.

[10] If we interpret zero utility of infinite disagreements in the DD game as $\lim_{t\to\infty} \delta^t$ with $\delta < 1$, then the corresponding utility of infinite disagreements in the DP game would be the limit of $-\delta^t$ with $\delta > 1$, which is negative infinity. We do not need this strong penalty to derive our theoretical results.

[11] Section 6 offers an alternative explanation and more discussion on the utility of infinite disagreements.

Let h^t denote the history at round t, including the identities of the previous proposers and the current proposer. Let $\{p_i^t(h^t), x_i^t(h^t)\}$ denote a feasible action for player i in round t, where $p_i^t(h^t) \in \Delta(P)$ is the (possibly mixed) proposal offered by player i as the proposer in round t, and $x_i^t(h^t)$ is the voting decision threshold of player i as a non-proposer in round t, where $\Delta(P)$ is the set of probability distributions of P. A strategy s_i is a sequence of actions $\{p_i^t(h^t), x_i^t(h^t)\}_{t=1}^{\infty}$, and a strategy profile s is an n-tuple of strategies, one for each player.

Concerning the DD game, it is known that there are numerous stage-dominated equilibria (Baron and Kalai, 1993), and virtually all allocations can be supported as an equilibrium under majority rule (Baron and Ferejohn, 1989). A similar folk theorem can be applied to the DP game.

Proposition 1. *Assume $n \geq q+1 \geq 3$ and $\delta \geq 1$. For any $p \in P$, there exists a stage-dominated equilibrium for which p is the equilibrium outcome.*

Proof: See Appendix A.

The result of Proposition 1 delivers a rationale for considering a refinement of the equilibria. We here focus on stationary subgame-perfect equilibria. A strategy profile is *stationary* if it consists of time- and history-independent strategies. A strategy profile is subgame perfect if no single deviation in a subgame can make the player better off.[12] Strategy s_i is now simplified to $\{p_i, x_i\}$. Furthermore, we consider symmetric agents, so the strategy boils down to (1) the proposal p when a member is recognized as a proposer and (2) the voting decision threshold x at which a non-proposer accepts. We also restrict our focus to equilibria in which each player's strategy is symmetric.[13]

3 Analysis

While the DD game has a unique SSPE in payoffs (Eraslan, 2002), the DP game has a continuum of stationary equilibria that involve different payoffs. For a brief illustration (only for Propositions 2 and 3), we start with a particular case in which a simple majority rule is applied, and $\delta = 1$. Perhaps the most intuitive stationary equilibrium involves allocation of the whole penalty to only one member.

Proposition 2 (Utmost inequality equilibrium)**.** *One SSPE can be described by the following strategy profile:*

[12] It is worth noting that a stationary equilibrium where everyone rejects every proposal forever is not subgame perfect: If one is offered a loss smaller than the ex-ante expected loss moving on to the next round, then deviating from the current "reject everything" strategy is at least weakly beneficial. This stationary (but not subgame-perfect) equilibrium, however, is more relevant in the cases where $\delta < 1$.

[13] If strategies, especially the coalition formation part of it, are asymmetric, then one can construct other non-stationary equilibria. For example, in three-player bargaining, if player 1 and player 2 always pick each other as a coalition partner, and player 3 picks one randomly, then the continuation value would differ. In general, if we allow any asymmetric mixing strategies in forming a minimum winning coalition, there will be a continuum of equilibria (Norman, 2002). We claim that this asymmetric type of equilibrium cannot be a proper ground for the experiment where subjects are randomly re-matched and their ID numbers are reassigned in every period.

- *Member i, being recognized as a proposer in round t, picks member $j \neq i$ at random and proposes $p_j = -1$ and $p_{-j} = 0$.*

- *A member offered no penalty accepts the proposal and rejects it otherwise.*

In this equilibrium, the proposal made by the first-round proposer is approved.

Proof: See Appendix A.

We call this equilibrium the utmost inequality (UI) equilibrium because only one member will be given the total burden of the penalty. Another equilibrium is the most egalitarian among stationary subgame-perfect equilibria.

Proposition 3 (Most egalitarian equilibrium)**.** *One SSPE can be described by the following strategy profile:*

- *The member recognized as the proposer in round t picks $\frac{n-1}{2}$ MWC members at random. She proposes $p_i = -1/n$ if $i \in MWC$, $s_{-i} = -\frac{n+1}{n(n-1)}$ if $i \notin MWC$ and keeps 0 for herself.*

- *If member i is offered $x \geq -1/n$, he accepts the proposal and rejects it otherwise.*

In this equilibrium, the proposal made by the first-round proposer is approved.

Proof: See Appendix A.

In the most egalitarian (ME) equilibrium, the distribution of the penalty is spread across members. Note that the ME equilibrium does not involve an equal split of the penalty: The allocation in the ME equilibrium is the most egalitarian in the sense that the largest share of the penalty that any one member would take is the smallest among all possible stationary equilibria.

Table 1 juxtaposes how the theoretical predictions of the DP game are different from those of the DD game under a simple majority rule when the discount factor is 1.

Table 1: Comparisons: Simple Majority, $\delta = 1$

Game	Proposer Share	MWC Share	non-MWC Share	Proposer Advantage†
DD	$1-\frac{n-1}{2n}$	$\frac{1}{n}$	0	$\frac{n-1}{2n}$
DP (UI)	0	0	1 (one of them)	0
DP (ME)	0	$\frac{1}{n}$	$\frac{n+1}{n(n-1)}$	$\frac{1}{n}$

†: Proposer advantage is a difference between the payoff of the proposer and that of the MWC member.

Indeed, there are other stationary subgame-perfect equilibria that take an intermediate form between the UI equilibrium and the ME equilibrium. For example, in one equilibrium, the proposer picks $\frac{n-1}{2}$ members randomly and offers $-\frac{2}{n-1}$ to each. The other $\frac{n-1}{2}$ members who were offered no penalty will accept the proposal. Propositions 4 and 5 describe all possible stationary subgame-perfect equilibria in the DP game for any $\delta \geq 1$.

Proposition 4. *Assume $q < n$. Every SSPE can be described by the following strategy profile:*

- *Member i, being recognized as the proposer in round t, selects $q-1$ MWC members at random. She proposes $p_j \geq -\delta/n$ if $j \in MWC$, proposes $p_j \leq 0$ if $j \in OTH \equiv N \setminus MWC \setminus \{i\}$ such that $\sum_{k \in OTH} p_k = -1 - \sum_{j \in MWC} p_j$, and keeps zero for herself.*
- *If member i is offered $x \geq -\delta/n$, he accepts the proposal and rejects it otherwise.*

In this equilibrium, the proposal made by the first-round proposer is approved.

Proof: See Appendix A.

The ME equilibrium and the UI equilibrium can be described as special cases of SSP equilibria constructed in Proposition 4. It is easy to characterize an equilibrium with respect to the payoff of an MWC member. The second bullet point of Proposition 4 specifies the maximum amount of loss willing to accept, $-\delta/n$, so a proposer should offer as low as $-\delta/n$ to the MWC members. Any offer between 0 and $-\delta/n$ to every MWC member is feasible, while it still allows the proposer to enjoy zero losses if $q < n$, as there is always at least one member who receives the remainders. The ME equilibrium arises when the MWC members are offered $-\delta/n$, while the UI equilibrium is when the MWC members are offered 0. If MWC members are offered an intermediate value between $-\delta/n$ and 0, then it constructs another stationary subgame-perfect equilibrium between the UI and the ME equilibria.

Proposition 5 describes the unique SSPE under a unanimity rule. From the perspective of the MWC members, the decision rule remains the same as in other voting rules: accept a proposal if it offers $x \geq -\delta/n$. Under a unanimity rule, however, the proposer's desire to minimize her own loss share drives to offer the MWC members the maximum willingness to suffer, $x = -\delta/n$.

Proposition 5. *Assume $q = n$. If $\delta \geq \frac{n}{n-1}$, proposer i still keeps zero for herself and offers $p_j \geq -\delta/n$ for all $j \neq i$. If $\delta < \frac{n}{n-1}$, the unique stationary subgame-perfect equilibrium is to offer $-\delta/n$ to every member and keep $\frac{(n-1)\delta - n}{n}$.*

Proof: See Appendix A.

There are at least three points worth mentioning. First, the theoretical predictions of the DP game, although the structure of the game can be understood as a mirror image of the DD game, are not the flip side of the theoretical predictions of the DD game except for the particular case where $n = 3$ and

$\delta = 1$.[14] By construction, the ME equilibrium corresponds to the SSPE in the DD game: The members in the MWC are offered the smallest amount of surplus that is just sufficient for them to accept the offer in the DD game, while they are offered the largest amount of losses that is just acceptable for them to agree to the offer in the DP game. We request attention to this result because at a mere glance, theoretical results seem symmetrical.

Second, the existence of other SSP equilibria allows us to consider many fragile aspects of the ME equilibrium, which is the mirror image of the essentially unique SSPE in the DD game. The equilibrium is not strict in the sense that players will vote for the proposal with probability 1 when indifferent between accepting and rejecting it. Even if the proposer decides to offer a loss to the MWC members that is "ε-less" than the continuation value, each player's "ε" may not be common knowledge, so choosing the ME strategy may not guarantee approval of the proposal. The cognitive cost for each player to coordinate on the ME equilibrium is also high: It requires each player to exactly calculate the continuation value given that other members also use the same stationary strategy, which varies by the voting rule, the size of the group, and the discount factor. In other words, the ME equilibrium is less robust given strategic uncertainty. Another notable observation is that when $\delta \geq \frac{n}{2(q-1)}$, the continuation value, the amount offered to the MWC members, can be *smaller* than the ex-ante payoff of the other members.[15] That is, the MWC members can be treated worse than other members in the ME equilibrium. In such a situation, the definition of the "minimum" winning coalition itself becomes fragile, as all the members receive an offer more attractive than their continuation value. Thus, the ME equilibrium, although it corresponds more directly to the unique equilibrium outcome in the DD game, is fragile in that it requires a stronger assumption about voting behavior and a higher coordination cost.

Third, the existence of multiple stationary subgame-perfect equilibria gives rise to the equilibrium selection issue.[16] Both the UI equilibrium and the ME equilibrium are optimal from a utilitarian perspective. Given the same level of social efficiency, which strategy would the proposer choose? On the one hand, the proposer may want to choose the most egalitarian strategy because the ME equilibrium is better from a Rawlsian perspective. Moreover, the ME equilibrium corresponds to the essentially unique SSPE in the DD game, so we set our null hypotheses based on the ME equilibrium. On the other hand, there may be an incentive for her to choose the most unequal strategy. If the proposer is uncertain about how often other players will mistakenly make a wrong decision, she may want to secure strictly more votes than q so that her payoff is robust to other members' mistakes. For this purpose, she may want to allocate the penalty to the smallest number of players. Taking inequity aversion

[14]When $n = 3$ and $\delta = 1$, the ME equilibrium allocation of the DP game, where the proposer keeps 0, a coalition member receives $-1/3$, and the other member received $-2/3$, looks like a mirror image of the SSPE allocation of the DD game, where the proposer keeps 2/3, a coalition member receives 1/3, and the other member receives nothing.

[15]For example, consider the ME equilibrium when $n = 5$, $q = 3$, and $\delta = 1.5$. Each of the MWC members is offered -0.3, while each of the other members is offered -0.2 on average. Considering this, we set δ for our experiment such that $\delta < \frac{n}{2(q-1)}$ under a majority rule.

[16]In Section 6.4, we discuss more equilibrium selection arguments, including the quantal response equilibrium, trembling hand perfection, and some behavioral arguments.

(Fehr and Schmidt, 1999) into account does not help us refine the set of equilibria.[17] From the perspective of the members who are offered a zero penalty in the UI equilibrium, although accepting the offer brings the largest disutility from the advantageous inequity perspective, it involves the smallest disutility from the disadvantageous inequity perspective.[18] Our laboratory experiments will answer this open question.

4 Experimental Design and Procedure

4.1 Design and Hypotheses

We tailored laboratory experiments to examine how people behave to determine the distribution of losses, especially in terms of the choices of the winning coalition. The major treatment variables address the group size ($n \in \{3,5\}$) and the voting rule ($q = (n+1)/2$ or majority; $q = n$ or unanimity). We set the appreciation factor δ to 1.2. Table 2 presents our 2×2 treatment design. Each of those treatments is respectively called M3 (majority rule for a group of three), M5 (majority + five), U3 (unanimity rule for a group of three), and U5 (unanimity + five). M3 and M5 are collectively called the majority treatments, and U3 and U5 are called the unanimity treatments.

Table 2: Experimental Treatments

		Voting Rule	
		Majority	Unanimity
Group Size	3	***M3***	***U3***
	5	***M5***	***U5***

Figure 2 illustrates the theoretical predictions, which can be categorized as two qualitatively different types. The first type of predictions (Hypotheses 1 and 2) include those that do not depend on equilibrium selection. The second type (Hypothesis 3) is those that vary depending on which equilibrium is selected.

[17] Similarly, taking loss aversion (Kahneman and Tversky, 2013) into account does not significantly help to further refine the set of equilibria, as we do not know the reference point of the players. If the reference point is set to zero, the "gain domain" is never achieved, so loss aversion does not play a role. If the reference point is set to an equally split loss, it implies that the reference point changes over time, which has little support. If the reference point is set to the ex-ante expected utility in the first round, there is still a continuum of equilibria, and the set could be *larger* than what we have, depending on the loss aversion parameters. Loss aversion could encourage the coalition members (who fear the possibility of losing more in the next round) to accept the less attractive offer now.

[18] Montero (2007) showed that in the DD game, inequity aversion might increase the proposer's share in equilibrium, and the underlying intuition follows the same logic. From the perspective of the coalition member, the marginal disutility from the increased difference between the proposer's share and what he is offered may be smaller than the marginal utility from the decreased difference between what he is offered and what other non-MWC members receive (zero).

The first set of hypotheses regards proposer advantage and bargaining efficiency, which are true regardless of which equilibrium is played in all treatments.

Hypothesis 1 (Proposer Advantage and Full Efficiency)**.**

(a) The proposer receives the smallest loss in all treatments.

(b) The first-round proposals are approved in all treatments.

The second set of hypotheses address the distribution of the loss. The agreed-upon shares of the proposer may vary across different group sizes depending on the voting rule. First, given the majority rule, the proposer's agreed-upon share is always zero regardless of the group size. Second, for a given group size, the proposer's agreed-upon share is larger under the unanimity rule than under the majority rule. Third, given the unanimity rule, the proposer's agreed-upon share is larger when the group size is smaller.

Hypothesis 2 (Share of Loss)**.**

(a) In the majority treatment, the proposer keeps zero regardless of the group size.

(b) The proposer keeps a smaller loss in the majority treatment than in the unanimity treatment.

(c) The proposer keeps a larger share of the loss in the U3 treatment than in the U5 treatment.

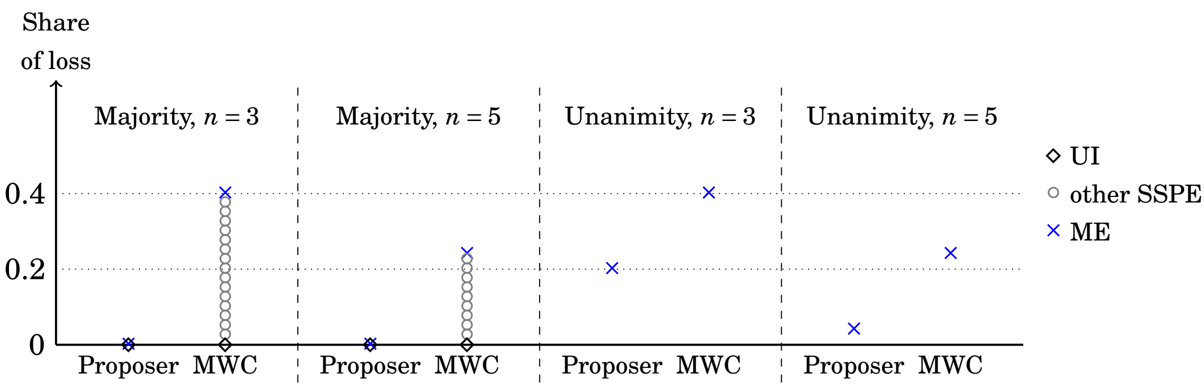


Figure 2: Hypotheses from Theoretical Predictions

This figure summarizes theoretical predictions corresponding to the experimental setup ($\delta = 1.2$). There is a continuum of other stationary subgame-perfect equilibria between the UI equilibrium and the ME equilibrium, mainly described by the share of the loss allocated to a minimum winning coalition member. The unique equilibrium under the unanimity rule corresponds to the ME equilibrium under the majority rule.

We move on to discuss the second type of predictions, which are dependent upon the equilibrium selection. First, the offers to the non-proposers vary based on the choice of an equilibrium. Especially in the majority treatment, players who are not the proposer are offered a share of the loss ranging from zero (to the MWC members in the UI equilibrium) to the full penalty (to the non-MWC members in the UI equilibrium). Second, under the majority rule, the range of shares that the MWC can receive in equilibrium varies by the group size, but such theoretical variations are not allowed under the unanimity rule. Given that our primary objective is to observe behaviors in the lab and falsify/select some equilibria, we shall derive our next set of *null* hypotheses based on the assumption that the ME equilibrium is played in the lab. We do not mean that we are selecting the ME equilibrium as the most plausible candidate; instead, it plays a role as the benchmark for clearly stating the experimental hypotheses. There are two reasons why we take the ME equilibrium as a benchmark. First, it is the closest to the mirror image of the unique stationary equilibrium of the DD game. Second, it is the unique stationary equilibrium prediction under the unanimity rule (i.e., when $q = n$). By continuity, it is natural to take the same equilibrium when $q < n$. In Figure 2, the upper bound of the MWC share of the loss (and the implied lower bound of the non-MWC share) constitute the ME equilibrium.

Hypothesis 3 (Winning Coalition and Non-proposers' Shares under Majority)**.**

In the majority treatments:

(a) The median share of the accepted proposal (or the largest share offered to the MWC members) is larger than the proposer's share.

(b) The agreed-upon share of the non-proposers who accept the proposal is larger in M3 than in M5.

(c) The number of non-proposers who accept the proposal is $(n-1)/2$. That is, one member rejects the proposal in the M3 treatment, and two members reject it in M5.

As we have emphasized already, the predictions summarized in Hypothesis 3 do not hold for the UI equilibrium. Contrary to Hypothesis 3 (a), the MWC members' shares are the same as that of the proposer in the UI equilibrium. In addition, the share of the accepting non-proposers is the same as zero in both majority treatments, so Hypothesis 3 (b) would be rejected under the UI equilibrium. While the ME equilibrium predicts that two members reject the proposal in the M5 treatment (Hypothesis 3 (c)), the UI equilibrium predicts that only one member will reject the proposal. Thus, testing Hypotheses 3 using the observed behaviors in the lab would enable us to justify one of the stationary equilibria. Given that the observed behaviors can be rationalized, we could conclude which equilibrium will be more likely to be selected.

4.2 Experimental Procedure

All the experimental sessions were conducted in English at the experimental laboratory of the Hong Kong University of Science and Technology in November 2018. The participants were drawn from the undergraduate population of the university. Four sessions were conducted for each treatment. A total of 271 subjects participated in one of the 16 ($= 4 \times 4$) sessions. Session sizes were 15 (in three M3, two M5, three U3, and one U5 sessions), 18 (in one M3 and one U3 sessions), or 20 (in two M5 and three

U5 sessions). Python and its application Pygame were used to computerize the games and to establish a server-client platform. After the subjects were randomly assigned to separate desks equipped with a computer interface, the instructor read aloud the instructions for the experiment. Subjects were also asked to carefully read the instructions, and then they took a quiz to demonstrate their understanding of the experiment. Those who failed the quiz were asked to reread the instructions and to retake the quiz until they passed. An instructor answered all questions until every participant thoroughly understood the experiment. Whenever a question was raised, the instructor repeated the question aloud and answered it so that every subject was equally informed.

We conducted many-person divide-the-penalty experiments. In structure, the game is a mirror image of a typical many-person divide-the-dollar game, and it proceeds as follows: At the beginning of each bargaining period (called a 'day' in the experiment), each bargainer is endowed with 400 tokens, a token being the currency unit used in the laboratory. In each bargaining round (called a 'meeting' in the experiment) one randomly selected player proposes a division of $-50*n$ tokens, where n is the number of players in each group. The proposal is immediately voted on. If the proposal gets q or more votes, the bargaining period ends, and the subjects' endowment is reduced based on the approved proposal. Otherwise, the bargaining proceeds to the second round, where the penalty increases by 20 percent, that is, in the second round, the players must determine an allocation of $-60*n$ tokens. A new proposer is randomly selected, and the new proposal is voted on. This process is repeated indefinitely until a proposal is passed.

Since the subjects were informed that they would eventually earn at least a show-up payment of HKD 30 ($\approx$ USD 4), we implicitly limited the largest possible losses out of the equilibrium. As long as the largest out-of-equilibrium loss is sufficiently large, in particular, as long as it is larger than $\delta*50*n$, no stationary equilibrium is restricted or ruled out. Thus, the theoretical analysis still serves as a benchmark for our experiments.[19]

Subjects in the U3 and M3 treatments participated in 12 bargaining periods and those in the U5 and M5 treatments participated in 15 bargaining periods.[20] We used the random matching protocol. Although new groups were formed every bargaining period, there was no physical reallocation of the subjects, and they only knew that they were randomly shuffled. They were not allowed to communicate with other participants during the experiment, nor were they allowed to look around the room. It was also emphasized to participants that their allocation decisions would be anonymous. Every sub-

[19] In a few cases under the unanimity treatment, the bargaining rounds went beyond the point where the total value of the loss exceeded the sum of the group members' show-up payments, but there was no noticeable discrepancy around the threshold round. Those periods were not selected as a paid period, so all subjects were paid strictly more than their show-up payment.

[20] The number of bargaining periods varies to increase the chances for every participant to play the proposer role. If there are 12 bargaining periods in treatments with $n = 5$, each subject could be recognized as a proposer 2.4 times on average, which is not large enough to observe variations by individual. We ex-post checked that every subject played the proposer role at least twice. We did not use the strategy method (i.e., asking all subjects to submit their proposals, knowing that one of them would be randomly selected for voting afterward) because we were unsure whether the strategy method, in this particular context of the DP game, would work the same as the standard method. Brandts and Charness (2011) report that 15 out of 29 existing comparisons between the two methods show either significant differences or some mixed evidence.

ject participated in only one of the sessions. The experimental instructions for the M5 treatment are presented in Appendix B.

At the end of the experiment, the subjects were asked to fill out a survey asking their gender and age as well as their degree of familiarity with the experiment. The subjects' risk preferences were also measured by the dynamically optimized sequential experimentation (DOSE) method (Wang et al., 2010). The number of tokens that each subject earned at one randomly selected period (Azrieli et al., 2018) was converted into HKD at the rate of 2 tokens = 1 HKD. The average payment was HKD 202.7 (≈ USD 26), including the HKD 30 guaranteed show-up fee. The payments were made in private, and subjects were asked not to share their payment information. Each session lasted 1.5 hours on average.

5 Experimental Results

This section presents our main findings by combining some test results for the hypotheses posed in the previous section. To recapitulate the types of the hypotheses, Hypotheses 1 and 2 are about the theoretical predictions commonly shared by any equilibrium in the continuum of the stationary subgame-perfect equilibria, and Hypothesis 3 regards theoretical predictions based on the ME equilibrium.

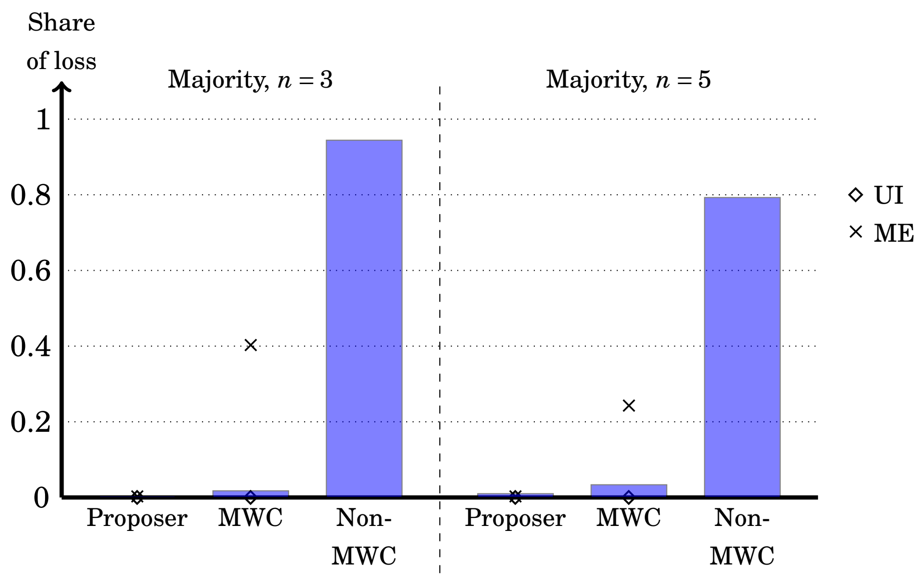


Figure 3: Proposed Shares, Majority
Approved proposals in the last 5 Periods

Figure 3, which juxtaposes the equilibrium predictions for the majority treatments and the observed average allocation of the loss from the approved proposals in the last five periods, represents the main findings. The share of the loss in the UI equilibrium is marked with ⋄, and that in the ME equilibrium is marked with ×. As is clearly illustrated, the observations from both the M3 and the M5 treatments reject the ME equilibrium predictions while supporting the UI equilibrium.

We reject[21] Hypothesis 3 (a) based on the ME equilibrium, as the MWC members in both the M3 and M5 treatments are offered almost zero shares of the loss. In the M3 treatment, the average share of the loss to the MWC member[22] in the last five periods is 0.0146 (2.19 tokens out of 150) larger than that to the proposer, but the difference is statistically insignificant (t-test, $p = 0.0782$),[23] and the magnitude is insubstantial. In the M5 treatment, the average share of the loss to the MWC members in the last five periods is 0.0114 (2.85 tokens out of 250) larger than that to the proposer, which is also statistically insignificant (t-test, $p = 0.3557$).

The average MWC share in the last five periods of the M5 treatment is 0.0150 larger than that of the M3 treatment, and the difference is weakly significant (t-test, $p = 0.0558$). This result strongly rejects Hypothesis 3 (b) based on the ME equilibrium in which the MWC member in the M3 treatment receives a larger share than the MWC members in the M5 treatment (t-test, one-sided, $p = 0.9721$).

When it comes to the loss share allocated to the non-MWC members, the rejection of the ME equilibrium predictions is even more evident. In the M3 treatment, one (non-MWC) member is offered almost the total loss, and such allocation is distinctively different from 0.6, the ME equilibrium share of the loss to the non-MWC member (t-test, $p < 0.0001$). The share of the non-MWC members is significantly larger than the ME equilibrium share of the loss (t-test, $p < 0.0001$).[24] In the M5 treatment, 61.90% of the approved proposals allocate almost the entire loss to one member, as only one member is offered a share of the loss larger than the proposal's median share. A total of 27.14% of the approved proposals distribute the almost entire loss to two members, as a share of the loss that is larger than the proposal's median share (which is near zero) is distributed to two members.[25] The average non-MWC share of the loss is approximately 0.7394, and the average number of the non-MWC members is 1.16, which leads us to reject Hypothesis 3 (c), as it is significantly smaller than 2 (t-test, $p < 0.0001$).

In both the M3 and M5 treatments, the proposer virtually keeps nothing for herself, consistent with the theoretical prediction (Hypothesis 2 (a)). In the M3 treatment, the proposer keeps, on average, 0.004 shares of the loss (0.6 tokens out of 150) for herself in the last five periods, which is insignificantly different from zero (t-test, $p = 0.3273$). In the last five periods of the M5 treatment, the proposer keeps, on average, 0.0212 shares of the loss (5.3 tokens out of 250) for herself. Although the average share of the loss is statistically different from zero (t-test, $p = 0.0435$), the magnitude is not substantial. In

[21]It may be more appropriate to say that we cannot reject the alternative hypothesis of Hypothesis 3 (a) rather than rejecting the null hypothesis itself.

[22]The MWC members are defined as the non-proposers who receive a share of the loss that is less than or equal to the proposal's median share. For example, in the M5 treatment, if a proposer offers $\{0, 0, 0, 0.4, 0.6\}$, then the median is 0 and the two members who received 0 are defined as the MWC. In the M3 treatment, a member receiving 0.3 from a proposal $\{0, 0.3, 0.7\}$ is defined as the MWC.

[23]Unless otherwise noted, the standard errors for the reported p-values of test statistics are clustered at the individual level. Using data from the last five periods allows us to give more weight to converged behavior, but the qualitative aspects of our findings remain unchanged if we use, for example, data from the last eight or from all periods.

[24]Alternatively, we also test if the average loss share of the member who rejects the proposal is the same as that in the ME equilibrium. In both the M3 and M5 treatments, test results confirm that the difference is significant at $p < 0.0001$.

[25]Approximately 11% of the approved proposals feature a grand coalition, where every member receives a share less than or equal to the proposal's median share of the loss. The proportion of such a grand-coalition proposal gradually decreases to 7% in the later periods.

addition, 60 out of the 70 approved proposals in the last five periods involve zero losses to the proposers.

Altogether, including the rejection of Hypothesis 3 and confirmation of Hypothesis 2 (a), our main finding from the majority treatments is the stark rejection of the ME equilibrium and the strong support of the UI equilibrium.[26]

Result 1. *In the majority treatments, experimental evidence rejects the ME equilibrium and supports the UI equilibrium.*

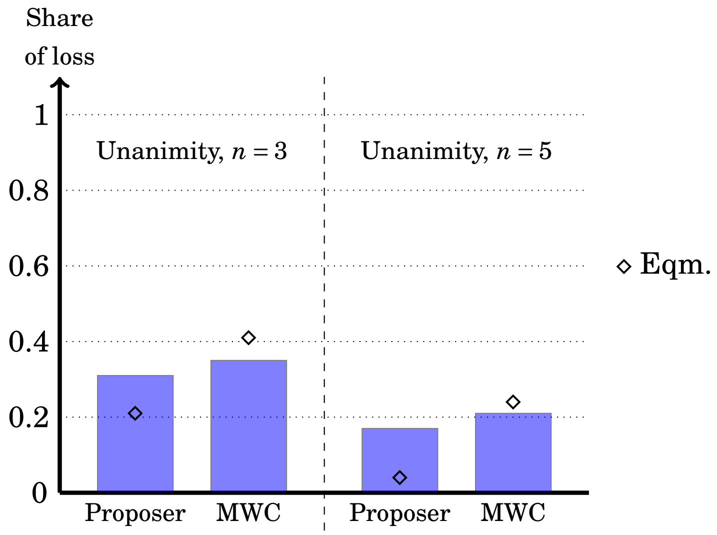
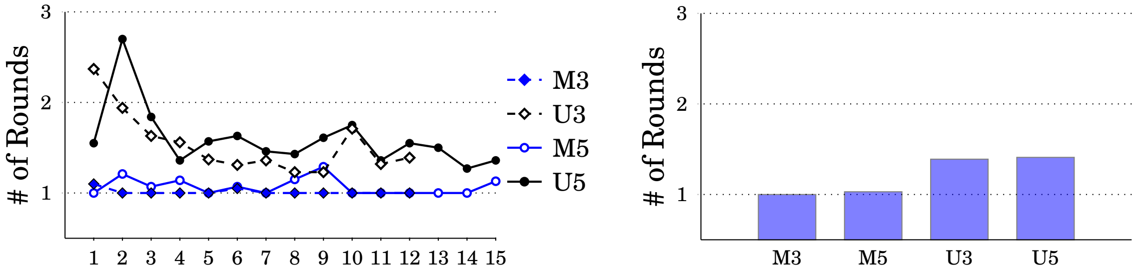


Figure 4: Proposed Shares, Unanimity
Approved proposals in the last 5 Periods

In the unanimity treatments, the allocation in the approved proposals is weakly consistent with the unique SSPE predictions. Figure 4 shows the theoretical predictions and the average share of the loss in the last five periods. The proposer's share is, on average, larger than the equilibrium level in both unanimity treatments. At the same time, the proposers offer a share of the loss to non-proposers that is smaller than the equilibrium level. The observation that the proposer keeps a significantly larger loss in the unanimity treatment than in the majority treatment is consistent with Hypothesis 2 (b). In addition, the proposer keeps a significantly smaller share of the loss than the other members in both unanimity treatments (t-tests, $p = 0.0185$ in U3 and $p < 0.0001$ in U5). Together with the majority treatments, we find that the observations are consistent with Hypothesis 1 (a). The non-proposers in the U3 treatment are offered a 0.1396 larger share of the loss than in the U5 treatment, and the difference is statistically significant ($p < 0.0001$), which confirms Hypothesis 2 (c). From the confirmations of Hypotheses 1 (a), 2 (b), and 2 (c), we draw our second result.

Result 2. *In the unanimity treatments, the allocations in the approved proposals are consistent with the theoretical predictions.*

[26]We discuss the selection of the UI equilibrium in more detail in Section 6.

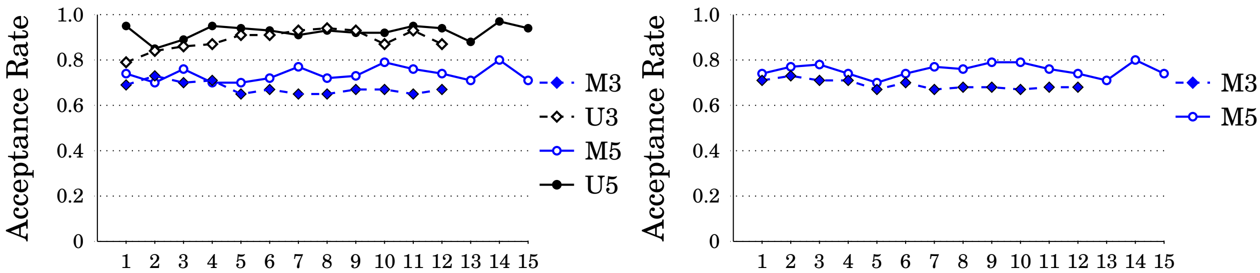


(a) Time Trend (b) Last 5 Periods

Figure 5: Average Number of Rounds

Figure 5 shows the average number of rounds by period. In the majority treatments, nearly all of the first-round proposals are approved, which is consistent with a theoretical prediction (Hypothesis 1 (b)). Even in the unanimity treatments, although the first three periods are somewhat varied (Figure 5 (a)), the average number of rounds in the last five periods is fewer than 1.5 (Figure 5 (b)). Albeit small, efficiency loss under a unanimity rule is one of the common findings in the multilateral bargaining experiments, such as Kagel et al. (2010), Miller and Vanberg (2013), and Kim (2018), to name a few.

Result 3. *The vast majority of the proposals are approved in the first round.*



(a) All Proposals (b) Approved Proposals

Figure 6: Average Acceptance Rate

Figure 6 shows the average proportion of subjects who accept a given proposal. In the M3 treatment, two-thirds of the subjects (two out of the three members) accept the proposal, which is consistent with the theoretical prediction for the size of the winning coalition. However, in the M5 treatment, nearly 80% of the subjects (four out of the five members) vote for the proposal. This observation is another way of rejecting Hypothesis 3 (a) in which two members reject the proposal.

Result 4. *In the majority treatments, the proposers form the winning coalition to minimize their losses. The size of the winning coalition is not necessarily minimal.*

Table 3 reports some regression results to examine whether individual characteristics affect the outcomes of the experiments. The regressions reported in Table 3 use only proposals that were approved in the first round. To summarize, we did not find any strong impact from individual characteristics. The dependent variable in the first three regressions is the proposer's own share, and the

Table 3: Individual Characteristics

Dep.Var.	Proposer's Own Share			Non-proposer's Vote		StDev(Proposal)
	(1)	(2)	(3)	(4: LPM)	(5: Logit)	(6)
M5	0.0170	0.0174*	0.0159	−0.0834***	−1.6503***	0.3182
	(0.0103)	(0.0100)	(0.0100)	(0.0222)	(0.4879)	(4.6753)
U3	0.2768***	0.2720***	0.2720***			−74.1292***
	(0.0116)	(0.0091)	(0.0092)			(1.9761)
U5	0.1522***	0.1523***	0.1503***			−76.7516***
	(0.0073)	(0.0071)	(0.0073)			(1.7611)
Share				−0.9659***	−7.7228***	
				(0.0249)	(0.7866)	
StDev				0.0006	0.0077	
				(0.0004)	(0.0047)	
Period	−0.0035***	−0.0039***	−0.0039***	−0.0019	−0.0186	1.3230***
	(0.0008)	(0.0008)	(0.0008)	(0.0022)	(0.0263)	(0.2992)
Female		0.0166**	0.0152**	−0.0171	−0.2057	−8.8552***
		(0.0067)	(0.0068)	(0.0273)	(0.3802)	(2.6477)
Age			−0.0063	0.0535**	0.7272*	3.6051
			(0.0088)	(0.0262)	(0.3712)	(4.0365)
RiskAversion			0.0020	0.0066	0.0894	−0.4723
			(0.0018)	(0.0063)	(0.0802)	(0.7085)
Familiarity			−0.0001	−0.0556	−0.9365	0.0337
			(0.0072)	(0.0347)	(0.4394)	(2.8944)
QuizFailed			−0.0050	0.0209	0.4068	0.6037
			(0.0071)	(0.0305)	(0.4615)	(2.6803)
_Cons.	0.0478***	0.0425***	0.0440***	0.9039***	3.0777***	73.0042***
	(0.0108)	(0.0099)	(0.0150)	(0.0555)	(0.7724)	(5.2324)
R^2	0.7058	0.7392	0.7434	0.6875	0.6360	0.7353
N	781	735	728	1239	1239	728

Only approved proposals in Round 1 are considered. In parentheses are standard errors clustered at the individual level. *, **, and *** indicate statistical significance at the 10% level, 5% level, and 1% level, respectively.

dependent variable in the last two regressions is the non-proposer's voting decision. Some explanatory variables are from the post-experiment survey. We collected self-reported gender and age. The subjects' risk preferences were measured by at most two survey questions, where the second question is dynamically adjusted based on the answer to the first question, which asks the subject to compare a simple lottery with a certain payment. This method enables us to categorize a subject into one of seven types of risk preferences. Familiarity is a subjective assessment of how familiar the subject was with the underlying game in the experiment. QuizFailed is a dummy variable indicating whether the subject had to retake the quiz after failing to pass, which would serve as a proxy for how well subjects understood the experiment. As control variables, we include treatment dummies and a time trend (labeled as Period) for regressions on the proposer's own share. We also include the offered share and the standard deviation of the proposal for regressions on the non-proposer's voting decision. The standard deviation of the proposal is added to examine whether the proposal's distributional shape matters in the subject's vote.[27] In all regressions, M3 is set as the baseline treatment. We focus on the approved proposals in round 1 only. Since the individual choices are positively correlated across periods, standard errors are clustered at the individual level.[28]

Risk preferences, familiarity, and comprehensibility of the experiment did not significantly affect the proposer's decisions or the non-proposer's voting. We found that females allocate slightly more (approximately 1.52% to 1.66%, varying by the model specification) losses to themselves. Related to this observation, we also found that female proposers are more egalitarian in how they allocate the losses: The standard deviation of the proposals offered by female proposers is, on average, 8.85 tokens smaller. Older subjects tended to accept the proposals more often, but the statistical significance is weak, and the age variance is not large, similar to many typical laboratory experiments.

Result 5. *Risk preferences, familiarity with the game, and comprehensibility were not significant factors affecting the outcomes of the experiments. Females tend to take a slightly greater share of the loss than males, and older subjects tend to accept the proposal.*

In summary, our experimental results are primarily consistent with the theoretical predictions based on the UI equilibrium, and individual characteristics do not lead to noticeably different experimental outcomes.

6 Discussions

In this section, we discuss some theoretical deviations to which we have paid less attention.

[27]For example, consider two proposals (0.2, 0, 0.8) and (0.2, 0.4, 0.4). For member 1, these two proposals offer the same amount of losses, 0.2, but the proposal's distribution varies. The standard deviation of the proposal will capture the impact of the distribution of the proposal if the subjects' voting decision is indeed affected by it.

[28]The standard errors clustered at the session level were overall smaller (that is, less conservative) than those at the individual level, so more estimated coefficients appear to be significant at the session level. Unless a subject randomly changes strategies over time, the individual choices are more correlated than the whole observation at the session level. Here, we report more conservative standard errors.

6.1 How to make the DD and the DP equivalent

We have claimed that the DD and DP games are fundamentally different, especially in terms of the number of SSP equilibria. Examining the conditions for those two games equivalent (except for the flipped signs and level shifts of the outcomes) would be worthwhile for illustrating the differences in another way. The difference between a growing penalty ($\delta \geq 1$) in the DP game and a discounting surplus ($\delta \leq 1$) in the DD game does not play an important role in leading to different theoretical results, as the theoretical differences remain unaffected when $\delta = 1$. Under a q-quota voting rule with $q < n$, there are two ways to make the theoretical predictions between the DD and the DP games similar.

1. In the DP game, if a maximum loss-share assigned to one player is capped at $\frac{n-\delta(q-1)}{n(n-q)}$, then only the ME remains as the SSPE.

2. In the DD game, if a maximum surplus share *to the proposer* (not every player) is capped at $\frac{\delta}{n}$, then there is a continuum of SSP equilibria where the equilibrium payoff of the MWC member varies from $\frac{\delta}{n}$ to $\frac{n-\delta}{n(q-1)}$.

If we define a "free surplus"[29] as the difference between the equilibrium payoff and the continuation value of one MWC member, then the uniqueness (in payoff) of the equilibrium in the DD game is described by the zero difference between the payoff and the continuation value. In other words, in the DD game, there is no free surplus that the proposer can add to the others. Meanwhile, in the DP game, the continuation value of the MWC member is $-\frac{\delta}{n}$, but the proposer can make a strictly better offer than $-\frac{\delta}{n}$ to his/her MWC members. That is, in the DP game, the natural cap of the proposer (who cannot enjoy more than zero losses) allows a free surplus to the MWC members. Thus, theoretical equivalence between the DD game and the DP game can be achieved either (1) if we restrict the proposer's free surplus in the DP game or (2) if we enforce that the proposer in the DD game to have a free surplus.[30]

6.2 Voting Rules Other Than Unanimity

Another issue may be the choice of voting rules other than unanimity. Since the UI equilibrium involves an extreme allocation of the loss to a few members, some risk-averse agents may demand unanimity. However, unanimity is not suitable for every situation. Implementation of a new policy would be one important example where a majority rule is applied. For example, the Tax Cuts and Jobs Act of 2017 in the United States was passed by the Senate on December 20, 2017, in a 51–48 vote. Assume for simplicity that a government wants to reform tax policy to cope with a budget deficit, and there are only three types of citizens with equal populations: the rich, the poor, and the middle-class.

[29] We thank an anonymous reviewer for the suggestion of this term.

[30] Although examining the conditions for making two games similar is theoretically entertaining, we are unsure whether we could find meaningful implications of this exercise. For example, to make the DP game be like the DD game, we should impose $\frac{n-\delta(q-1)}{n(n-q)}$ as a loss cap. When we ask ourselves about the use of this result or about the practical resemblance in distributive politics, it would not seem to add much value.

In this case, a policy victimizing one of the three distinct groups by allocating the tax burden to that group may be implemented, but we do not claim that we should change the voting rule to unanimity due to that possibility. In addition, although the stability of the voting rule is beyond our concerns in this paper, studies including Barbera and Jackson (2004) characterize a self-stable majority voting rule with the persuasive argument that the general trend is away from unanimity. Moreover, as our experimental evidence and many other similar experimental studies show, a unanimity rule leads to efficiency loss due to delay.[31] Risk-neutral agents who negotiate over a loss repeatedly may want to avoid unanimity because it might eventually be harmful to every agent.

6.3 Bargaining When Delay is Socially Desirable

We assume $\delta = \beta g \geq 1$ so that no one has an incentive to postpone their bargaining decision. However, in situations where $\delta < 1$, that is, β (the subjective discount factor of a future payoff) is sufficiently smaller than $1/g$ (the inverse of the growth rate of the penalty), the Pareto-optimal allocation is for everyone to reject any form of proposal for any round t so that everyone will eventually have zero losses. In this situation, still, the stationary subgame-perfect equilibria can be sustained as long as we maintain the assumptions that each individual is self-interested and that subgame-perfect strategies are considered. For example, when a proposal allocating all losses to one member is put to a vote, a member who receives an offer of zero losses would accept the proposal because the continuation value of the next bargaining round is at least weakly smaller than the zero losses. If the qualified number of votes for approval is less than n, the proposal would be accepted immediately. Similar to the public goods game situation, the Pareto-optimal collective behavior is distinctly different from the equilibrium behavior.

We have paid less attention to the case with $\delta < 1$ for several reasons. First, we have tried to make the structure of the DP game as similar to that of the DD game as possible. In the DD game, delay is discouraged, as it is in the DP game with $\delta \geq 1$. Second, the experimental evidence may be confounded because each subject's internalized social norms may be heterogeneous and unobservable (Kimbrough and Vostroknutov, 2016). If the primary purpose of this study has been to observe how subjects behave differently when the Pareto-optimal behavior and the equilibrium behavior diverge, a typical linear public goods game would have been more pertinent. Third, since it is unusual to have losses that will disappear as time passes if nothing is done, we claim that $\delta < 1$ is less relevant to real-life situations.

6.4 Regarding the Utility of Infinite Disagreements

In the DD game, the utility when disagreeing forever is assumed to be zero, which is the lowest possible earning in the first round. To have a corresponding form, we assume that the utility of infinite disagreements is negative one, which is the largest possible loss in the first round. However, the

[31] Bouton et al. (2018) discourage the use of a unanimity rule for a different reason: by showing that unanimity is Pareto-inferior to majority rules with veto power.

purpose of setting it to -1 is merely to have a corresponding form. As long as the utility is smaller than the continuation value of the second round, no stationary subgame-perfect equilibria predict that players will move to the second round. As we explained in the previous subsection, the equilibria are sustained even when the loss exponentially decreases to zero.

It may be easier to deal with the utility of infinite disagreements in the following alternative interpretation. Suppose that we still have the same growth rate $g \geq 1$ and the discount factor $\beta \leq 1$. Instead of assuming that utility is accrued either when an agreement is reached or when infinite disagreements occur, assume that each disagreement renders a disutility of $\frac{g-1}{n}$ (the equal split of the increased amount of loss) to all players. One can imagine that each delay in an environmental policy agreement leads everyone to gain a (marginal) disutility from the untreated environmental damage. Then, the utility of infinite disagreements is the sum of the geometric sequences, $\sum_{t=0}^{\infty} -\beta^t \frac{g-1}{n} = \frac{1-g}{n(1-\beta)}$. To characterize the stationary subgame-perfect equilibria in which disagreeing forever is not incentivized, it is sufficient to assume that $\frac{1-g}{n(1-\beta)}$ is smaller than the ex-ante expected payoff of bargaining, $-\frac{1}{n}$. In other words, the parametric assumption we have essentially in mind is $\frac{1-g}{n(1-\beta)} < -\frac{1}{n}$, or $g+\beta > 2$, and not exponentially increasing losses. This assumption holds with the entire set of parameter values we have considered except for the indeterministic case $g = \beta = 1$.

6.5 Equilibrium Selection

Our primary purposes are to convince that the DP game is theoretically different from the DD game and to report that the experimental observations are quite distinct. However, it is worth discussing what the proper refinement of the equilibrium of giving zero losses to all winning coalition members is. Indeed there are many justifications for selecting the UI equilibrium.

Although we did not explicitly mention the quantal response equilibrium (QRE, McKelvey and Palfrey, 1995), one argument for why the ME equilibrium is fragile goes along with the assumption of QRE. If the winning coalition members might sometimes mistakenly reject the proposal, then the proposer needs to minimize the risk associated with such mistakes by providing them with more favorable offers. QRE has a property in which the probability of a mistake depends on the cardinal payoff that a player gives up, so it renders the proper incentives to choose a particular proposal for which approval depends least on the critical calculation of the indifferent offer.

While the idea of QRE can address why each of winning coalition members could have the least losses, trembling hand perfection (THP, Selten, 1975) could explain why the size of the winning coalition could be larger than the minimum when it is possible. The possibility of nonproposers' mistakes will lead the proposer to demand a larger coalition. When $n = 5$, for example, among several stationary equilibria allocating zero losses to the winning coalition members, $(0,0,0,-x,-1+x)$, THP will select the most uneven allocation, $x = 0$, because this is the way to minimize the risk of rejection due to mistakes. This argument is consistent with our experimental findings in M5.

If we seek behavioral arguments, in-group favoritism can explain the selection of the UI equilibrium as well. In-group favoritism is one empirical similarity between our experimental observations and those in previous experiments involving the DD game (Fréchette et al., 2005). Gamson's Law, a popular empirical model that supports an equal split within a coalition, is often interpreted as evidence of in-group favoritism, which might lead to the proposer's partial rent extraction as opposed to the full rent extraction predicted by the Baron-Ferejohn model. In the sense that in the UI equilibrium, the proposer treats the MWC members most favorably, our observations might be consistent with the empirical interpretation of in-group favoritism. From the perspective of the proposer who has an epsilon concern regarding in-group favoritism, allocating zero to the MWC members is a corner solution, regardless of how negligible the in-group favoritism may be. Experimental evidence, including Efferson et al. (2008), suggest that in-group favoritism can evolve with arbitrary and initially meaningless markers. Although in our experimental setting, there are no clear distinctions between in-group and out-group, a sense that some members must vote "yes" for the proposer might be sufficient to form a notion of an in-group.

Last, if subjects are concerned about utilitarian social welfare and every subject has a concave utility on the losses, then the UI equilibrium is likely to be selected. If the marginal disutility of a loss is diminishing, as loss-averse utility functions are typically characterized, the disutility of one person's significant loss is smaller than the sum of disutilities of several persons' small losses. Then, selecting the UI equilibrium leads to the largest utilitarian social welfare. Although we believe those are plausible arguments, we admit that our experiments are not suitable for determining which arguments are more plausible than others.

7 Concluding Remarks

We examine the divide-the-penalty (DP) game to better understand multilateral bargaining when agents are dealing with the distribution of a loss. Although the literature on multilateral bargaining is substantial, both theoretically and experimentally, multilateral bargaining over the division of losses has received less attention. This may be, perhaps, because we have naïvely conjectured that the theoretical properties of the DP game are a mirror image of those from the divide-the-dollar (DD) game due to their structural resemblance. We theoretically show that there are fundamental differences. The stationary subgame-perfect equilibria in the DP game are no longer unique in payoffs, unlike the DD game. One extreme among the continuum of stationary subgame-perfect equilibria, which we call the most egalitarian (ME) equilibrium, is characterized similarly to the unique SSPE in the DD game. The other extreme equilibrium, which we call the utmost inequality (UI) equilibrium, predicts that the proposer concentrates the penalty on a few members. Although the ME equilibrium shares many properties with the SSPE in the DD game, experimental evidence is primarily consistent with the predictions based on the UI equilibrium.

Our results have at least two implications. First, multilateral bargaining over the division of losses should not be understood through the lens of the typical DD game because both the theoretical properties and the experimental evidence deviate from those of the DD game. Second, many interesting studies in multilateral bargaining on a gain domain are worth revisiting. Bargaining among asymmetric players, dynamic multilateral bargaining, the allocation of public bads produced for the agents' private gains, and changes in the bargaining protocols including competitions for recognition are some, but not all, subjects that can extend this study. The direction of research should distinguish simple behavioral/psychological framing effects from more fundamental differences.

References

Aghion, Philippe, Alberto Alesina, and Francesco Trebbi, "Endogenous Political Institutions," *The Quarterly Journal of Economics*, 05 2004, *119* (2), 565–611.

_ **and Patrick Bolton**, "Incomplete Social Contracts," *Journal of the European Economic Association*, 2003, *1* (1), 38–67.

Agranov, Marina and Chloe Tergiman, "Communication in multilateral bargaining," *Journal of Public Economics*, 2014, *118*, 75–85.

Ali, S. Nageeb, B. Douglas Bernheim, and Xiaochen Fan, "Predictability and Power in Legislative Bargaining," *Review of Economic Studies*, 2019, *86* (2), 500–525.

Andreoni, James, "Warm-Glow Versus Cold-Prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments," *The Quarterly Journal of Economics*, 1995, *110* (1), 1–21.

Azrieli, Yaron, Christopher P. Chambers, and Paul J. Healy, "Incentives in Experiments: A Theoretical Analysis," *Journal of Political Economy*, 2018, *126* (4), 1472–1503.

Baranski, Andrzej and John H. Kagel, "Communication in Legislative Bargaining," *Journal of the Economic Science Association*, Jul 2015, *1* (1), 59–71.

Barbera, Salvador and Matthew O. Jackson, "Choosing How to Choose: Self-Stable Majority Rules and Constitutions," *The Quarterly Journal of Economics*, 2004, *119* (3), 1011–1048.

Baron, David and Ehud Kalai, "The Simplest Equilibrium of a Majority-Rule Division Game," *Journal of Economic Theory*, 1993, *61* (2), 290–301.

Baron, David P. and John A. Ferejohn, "Bargaining in Legislatures," *American Political Science Review*, 1989, *83* (4), 1181–1206.

Battaglini, Marco and Stephen Coate, "Inefficiency in Legislative Policymaking: A Dynamic Analysis," *The American Economic Review*, 2007, *97* (1), 118–149.

Bernheim, B. Douglas, Antonio Rangel, and Luis Rayo, "The Power of the Last Word in Legislative Policy Making," *Econometrica*, 2006, *74* (5), 1161–1190.

Bogomolnaia, Anna, Hervé Moulin, Fedor Sandomirskiy, and Elena Yanovskaia, "Dividing bads under additive utilities," *Social Choice and Welfare*, Oct 2018.

Bouton, Laurent, Aniol Llorente-Saguer, and Frédéric Malherbe, "Get Rid of Unanimity Rule: The Superiority of Majority Rules with Veto Power," *Journal of Political Economy*, 2018, *126* (1), 107–149.

Brandts, Jordi and Gary Charness, "The strategy versus the direct-response method: a first survey of experimental comparisons," *Experimental Economics*, Sep 2011, *14* (3), 375–398.

Christiansen, Nels and John H. Kagel, "Reference point effects in legislative bargaining: experimental evidence," *Experimental Economics*, Sep 2019, *22* (3), 735–752.

_ , Tanushree Jhunjhunwala, and John Kagel, "Gains versus Costs in Legislative Bargaining," 2018. Working paper.

Diermeier, Daniel and Antonio Merlo, "Government Turnover in Parliamentary Democracies," *Journal of Economic Theory*, 2000, *94* (1), 46–79.

_ and Pohan Fong, "Legislative Bargaining with Reconsideration," *Quarterly Journal of Economics*, 2011, *126* (2), 947–985.

_ and Rebecca Morton, "Experiments in Majoritarian Bargaining," in "Studies in Choice and Welfare" Studies in Choice and Welfare, Springer Berlin Heidelberg, 2005, chapter Social Choice and Strategic Decisions, pp. 201–226.

Efferson, Charles, Rafael Lalive, and Ernst Fehr, "The Coevolution of Cultural Groups and Ingroup Favoritism," *Science*, 2008, *321* (5897), 1844–1849.

Eraslan, Hülya, "Uniqueness of Stationary Equilibrium Payoffs in the Baron-Ferejohn Model," *Journal of Economic Theory*, 2002, *103* (1), 11–30.

_ and Andrew McLennan, "Uniqueness of Stationary Equilibrium Payoffs in Coalitional Bargaining," *Journal of Economic Theory*, 2013, *148* (6), 2195–2222.

_ and Kirill S. Evdokimov, "Legislative and Multilateral Bargaining," *Annual Review of Economics*, 2019, *11*, 443–472.

Fehr, Ernst and Klaus M Schmidt, "A Theory of Fairness, Competition, and Cooperation," *Quarterly journal of Economics*, 1999, pp. 817–868.

Feinerman, Eli, Israel Finkelshtain, and Iddo Kan, "On a Political Solution to the NIMBY Conflict," *The American Economic Review*, 2004, *94* (1), 369–381.

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. Special Issue in Honor of Richard D. McKelvey.

Fréchette, Guillaume R., John H. Kagel, and Massimo Morelli, "Pork versus public goods: an experimental study of public good provision within a legislative bargaining framework," *Economic Theory*, 2012, *49* (3), 779–800.

Fréchette, Guillaume R., John H. Kagel, and Steven F. Lehrer, "Bargaining in Legislatures: An Experimental Investigation of Open versus Closed Amendment Rules," *American Political Science Review*, 2003, *97* (2), 221–232.

Fredriksson, Per G., "The Siting of Hazardous Waste Facilities in Federal Systems: The Political Economy of NIMBY," *Environmental and Resource Economics*, Jan 2000, *15* (1), 75–87.

Gaertner, Wulf, Richard Bradley, Yongsheng Xu, and Lars Schwettmann, "Against the proportionality principle: Experimental findings on bargaining over losses," *PLOS ONE*, 07 2019, *14* (7), 1–18.

Gerardi, Dino, Margaret A. McConnell, Julian Romero, and Leeat Yariv, "Get Out the (Costly) Vote: Institutional Design for Greater Participation," *Economic Inquiry*, 10 2016, *54* (4), 1963–1979.

Harstad, Bård, "Majority Rules and Incentives," *The Quarterly Journal of Economics*, 11 2005, *120* (4), 1535–1568.

_ , "Strategic Delegation and Voting Rules," *Journal of Public Economics*, 2010, *94* (1), 102–113.

Jackson, Matthew O. and Boaz Moselle, "Coalition and Party Formation in a Legislative Voting Game," *Journal of Economic Theory*, 2002, *103* (1), 49–87.

Kagel, John H., Hankyoung Sung, and Eyal Winter, "Veto Power in Committees: An Experimental Study," *Experimental Economics*, 2010, *13* (2), 167–188.

Kahneman, Daniel and Amos Tversky, "Prospect theory: An analysis of decision under risk," in "Handbook of the fundamentals of financial decision making: Part I," World Scientific, 2013, pp. 99–127.

Kim, Duk Gyoo, ""One Bite at the Apple": Legislative Bargaining without Replacement," 2018. Working paper.

_ , "Recognition without Replacement in Legislative Bargaining," *Games and Economic Behavior*, 2019, *118*, 161–175.

Kimbrough, Erik O. and Alexander Vostroknutov, "Norms Make Preferences Social," *Journal of the European Economic Association*, 2016, *14* (3), 608–638.

Levinson, Arik, "NIMBY taxes matter: the case of state hazardous waste disposal taxes," *Journal of Public Economics*, 1999, *74* (1), 31–51.

McKelvey, Richard D. and Thomas R. Palfrey, "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behavior*, 1995, *10* (1), 6–38.

Miller, Luis and Christoph Vanberg, "Decision Costs in Legislative Bargaining: An Experimental Analysis," *Public Choice*, 2013, *155* (3), 373–394.

Montero, Maria, "Inequity Aversion May Increase Inequity," *The Economic Journal*, 2007, *117* (519), C192–C204.

Norman, Peter, "Legislative Bargaining and Coalition Formation," *Journal of Economic Theory*, 2002, *102* (2), 322–353.

Ok, Efe A., "On the principle of equal sacrifice in income taxation," *Journal of Public Economics*, 1995, *58* (3), 453–467.

Palfrey, Thomas R., "Experiments in Political Economy," *Handbook of Experimental Economics*, 2016, *2*, 347–434.

Peterson, Elisha and Francis Edward Su, "Four-Person Envy-Free Chore Division," *Mathematics Magazine*, 2002, *75* (2), 117–122.

Selten, Reinhard, "Reexamination of the perfectness concept for equilibrium points in extensive games," *International Journal of Game Theory*, Mar 1975, *4* (1), 25–55.

Stromquist, Walter, "How to Cut a Cake Fairly," *The American Mathematical Monthly*, 1980, *87* (8), 640–644.

Volden, Craig and Alan E. Wiseman, "Bargaining in Legislatures over Particularistic and Collective Goods," *American Political Science Review*, 2007, *101* (1), 79–92.

Wang, Stephanie, Colin F. Camerer, and Michelle Filiba, "Dynamically Optimized Sequential Experimentation (DOSE) for Estimating Economic Preference Parameters," 2010. Mimeo.

Weg, Eythan and Rami Zwick, "On the robustness of perfect equilibrium in fixed cost sequential bargaining under an isomorphic transformation," *Economics Letters*, 1991, *36* (1), 21–24.

Young, H. Peyton, "Distributive justice in taxation," *Journal of Economic Theory*, 1988, *44* (2), 321–335.

A Appendix - Proofs

Proof of Proposition 1: This proof is analogous to the proof of Proposition 2 in Baron and Ferejohn (1989). Fix a strategy profile with the following statements.

1. For all $i \in N$ and $t \in \mathbb{N}_+$, if i is recognized in t, i proposes $p^{it} = p$, and all individuals vote for p.
2. If p is rejected under the q−quota rule in t, $j \in N$ recognized in $t+1$ proposes $p^{j(t+1)} = p$.
3. If, in any period t, the chosen proposer i offers an alternative to p, say $p^{it} = y \neq p$, then
   (3.a) a set $M(y)$ of at least q individuals rejects y;
   (3.b) the period $t+1$ proposer, say j, offers an allocation z such that $z_i = -1$ and all individuals in $M(y)$ vote for z against y.
4. If, in (3.b), the period $t+1$ proposer j offers some alternative $y' \neq z$, repeat (3) with y' replacing y and j replacing i.

Statement 1 specifies what happens along the equilibrium path. Statements 2, 3, and 4 describe off-the-equilibrium path behavior. That is, they jointly specify the consequences of any deviation from the behavior specified in 1.

For notational simplicity, relabel players such that player n is the period t proposer who offers $p^{nt} = y \neq p$, where $p_n < 0$, and $y_j \leq y_{j+1}$ for all $j = 1, \ldots, n-2$. If $p_n = 0$, there is no way for player n to be better off, so $p_n < 0$ is reasonable without loss of generality. It is trivial that $y_n > -1$, because player n does not have an incentive to deviate from p to keep the all the loss from the beginning. Under (3.a) and (3.b), players in $M(y)$ reject y, and the next proposer offers an alternative proposal z with $z_n = -1$ such that $M(y)$ approves. Under such a distribution z, (3.a) and (3.b) describe best response behavior to y. We divide situations into two cases: Assume first that the proposer conditional on y being rejected is some individual $j \neq n$. Let $M^*(y) = \{1, \ldots, q\}$, let $Y^* = \sum_{i \in M^*(y)} y_i$, and let $m^* = |\{i \in M^*(y) : y_i < 0\}|$. By construction of $M^*(y)$, $Y^* < 0$ and $m^* > 0$. $Y^* = 0$ (and hence $m^* = 0$) implies that $p_n = -1$ and thus $p = z$. If $y_i < 0$, then z is strictly preferred because $\delta z_i = 0 > y_i$. If $y_i = 0$, then z is as preferred as y because the payoff of i is unaffected. If z is rejected, then under strategy statement 4, it will simply become the next proposal and so on.

Now we assume that player n is again recognized as a proposer in the next period. Our goal is to show that player n cannot benefit from proposing any allocation other than p. In such a case, (3.b) specifies that player n proposes the allocation z, which "punishes" herself for her initial deviation. Should she fail to do this and instead propose some $y' \neq z$, strategy statement 4 requires a q-majority to reject y' and the period $t+2$ agenda-setter to offer z, which then passes. Therefore, the only circumstance under which the period t proposer n can avoid having z proposed and accepted in response to an initial deviation to $y \neq p$ is when player n is chosen in every period as the proposer. Such probability $(1/n)^t$ approaches zero, and the size of the penalty for the deviation is non-decreasing. Therefore, player n

is not better off by deviating from proposing p, as her hope that she could eventually attain a higher payoff than proposing p and accepting p_n is futile. $\square$

Proof of Proposition 2: Suppose for every round, players have an identical stationary strategy as described above. A member who received an offer of zero penalties this round will accept the proposal if moving on to the next round does not make him better off. In the next round, with probability $(n-1)/n$, he will be a proposer or a member who receives no penalty. With probability $1/n$, he will be randomly selected by a proposer in that round and take all the penalty. Certainly, if the utility from the current offer (zero) is strictly larger than the continuation value $(-\delta^{t-1})$, he will accept the offer. The proposer, who keeps no penalty for herself, cannot be better off by any other proposal. Thus, no one would be better off by deviating from this stationary strategy profile for any round. $\square$

Proof of Proposition 3: Consider player i who received an offer of $-1/n$. If the game moves on to the next round, his expected payoff is

$$\frac{1}{n}0-\frac{n-1}{2n}\frac{1}{n}-\frac{n-1}{2n}\frac{n+1}{n(n-1)}=-\frac{n-1}{2n^2}-\frac{n+1}{2n^2}=-\left(\frac{2n}{2n^2}\right)=-\frac{1}{n}.$$

Therefore, he will not be better off by rejecting the current offer. From the perspective of the current proposer, there is no strategy to make her better off than receiving zero penalties. $\square$

Proof of Proposition 4: First we show that under any voting rule except for unanimity, there is no stationary equilibrium where the proposer keeps a strictly negative payoff.

Lemma 1. *For any $q < n$, the proposer's share in the proposal of any of SSPE is zero.*

Proof: Without loss of generality, relabel member 1 as the proposer in the first round, and $p_i \geq p_{i+1}$ for $i = 1,\ldots,n-1$. Suppose for the contradiction that $p_1 < 0$. There could be at most $n-q$ members who vote against the proposal. Define $M(p)$ as a set of members who vote against the proposal. If $M(p)$ is nonempty, consider an alternative proposal p' that subtracts p_1 from p and adds p_n to one randomly selected member in $M(p)$. p' would make the proposer better off, while the members who vote for the proposal are not affected, because the continuation value under a stationary proposal p' is identical to that under p, that is,

$$\delta\sum_{i=1}^{n}\frac{1}{n}p_i=-\frac{\delta}{n}=\delta\frac{1}{n}\sum_{i=1}^{n}\frac{1}{n}p'_i.$$

Therefore, the proposer has an incentive to deviate from the equilibrium proposal, which contradicts the supposition of subgame perfection. $\square$

Next, for any stationary strategies, the continuation value is $-\frac{\delta}{n}$. Suppose that a proposer in the next round offers $(p_1, \ldots, p_n)$. For any player i, the expected payoff of moving on to the next round is:

$$\delta\left(\frac{1}{n}p_1 + \cdots + \frac{1}{n}p_n\right) = \frac{\delta}{n}\sum_{i=1}^{n} p_i = -\frac{\delta}{n}.$$

Therefore, players offered a share greater than $-\frac{\delta}{n}$ are willing to accept the current proposal. Since the proposer, who keeps zero (Lemma 1), wants her proposal to be approved, she must offer more than $-\frac{\delta}{n}$ to $q-1$ players. The allocation of the remaining losses, $-1-\sum_{j\in MWC} p_j$ must be allocated to the other members who are not included as a minimum winning coalition. □

Proof of Proposition 5: As long as players use a stationary strategy, the continuation value is $-\frac{\delta}{n}$. If the proposer offers $-\frac{\delta}{n}$ to every player, then $-1+\frac{(n-1)\delta}{n}$ is the remaining loss that she would take. If $\delta \geq \frac{n}{n-1}$, then $-1+\frac{(n-1)\delta}{n} > 0$. That is, the proposer still has a room to keep zero for herself, and allocate the losses unevenly to other players as long as the offer made to other players is greater than or equal to $-\frac{\delta}{n}$. If $\delta < \frac{n}{n-1}$, however, the proposer must keep $\frac{(n-1)\delta}{n} - 1$ for herself and offer $-\frac{\delta}{n}$ to the other members.

B Appendix - Experimental Instructions (M5)

Welcome to this experiment. Please read these instructions carefully. 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

The experiment consists of 15 "Days." In each Day, every participant will be endowed with 400 tokens, and you will be randomly matched with four other participants to form a group of five. The five group members need to decide how to split a DEDUCTION of (at least) 250 tokens from group members' endowments.

How the groups are formed

In each Day, all participants will be randomly assigned to groups of five members. Each member of a group is assigned an ID number (from 1 to 5), which will be displayed on the top of the screen. In a given Day, once your group is formed, the five group members will not change. Your ID is fixed throughout the Day.

Once the Day is over, you will be randomly re-assigned to a new group of five, and you will be assigned a new ID. Check your ID number when making your decisions.

You will not learn the identity of the participants you are matched with, nor will those participants learn your identity. Identities remain anonymous even after the end of the experiment.

How a deduction of tokens is divided

In each Day, you and your group members will decide how to split a deduction of (at least) 250 tokens across group members. Each Day may consist of several 'Meetings.'

In Meeting 1, one of the five members in your group will be randomly chosen to make a proposal to split the deduction of 250 tokens as follows.

	Member 1	Member 2	Member 3	Member 4	Member 5
# of Tokens Deducted:	____________	____________	____________	____________	____________

The number of tokens deducted from each member must be between 0 and 250. The total number of tokens must add up to 250 tokens.

Each member has the same chance of being chosen to be the proposer. After the proposer has made his/her proposal, the proposal will be voted up or down by all members of the group. Each member, including the proposer, has one and only one vote.

- If the proposal gets three or more votes, it is approved. The tokens allocated to you are DE-

DUCTED from your endowment and then the day ends.

- Otherwise, the proposal is rejected and your group moves to Meeting 2.

In Meeting 2, one member will be randomly selected to be a proposer. Every member, including the proposer in Meeting 1, has an equal chance to be a proposer. The total amount of tokens to be deducted will increase by 20% of that in the previous Meeting. That is, the five members in Meeting 2 need to decide how to split a deduction of 300 tokens. After the proposer proposes how to split the deduction of 300 tokens, it will be voted up or down by all members of the group. If this new proposal is rejected in Meeting 2, then in Meeting 3, another randomly selected member proposes to how to split a deduction of 360 tokens (20% more of 300 tokens), and so on. Your group will repeat the process until a proposal is approved. The following table shows the size of the deduction of tokens for each meeting.

Meeting	1	2	3	4	5	6	7	...
Deduction (in Tokens)	250	300	360	431	518	622	746	20% Larger

The amount of tokens you need to deduct is growing

To summarize, if you are selected as a proposer, make a proposal of splitting the deduction of the current number of tokens, and move to the voting stage. If you are not a proposer, wait until the proposer makes a proposal, examine it and decide whether to accept or reject it. Previous proposers can be a proposer again. If a proposal is approved, the number of tokens offered to you will be DEDUCTED from your endowment.

Information Feedback

At the end of each Meeting, you will be provided with a summary of what happened in the Meeting, including the proposed split of the deduction, the proposer's ID, and the voting outcome. At the end of each Day, you will learn the approved proposal and your earning from the Day.

Payment

In each Day, your earning is

[400 tokens − the number of tokens offered to you in the approved proposal]

The server computer will randomly select one Day and your earning in that Day will be paid. Each day has an equal chance to be selected for the final cash payment. So it is in your best interest to take each Day equally seriously. Your total cash payment at the end of the experiment will be the number of tokens you earn in the selected Day converted into HKD at the exchange rate of 2 tokens = 1 HKD plus 30 HKD guaranteed show-up fee.

Summary of the process

1. The experiment will consist of 15 Days. There may be several Meetings in each Day.
2. Prior to each Day, every participant is endowed with 400 tokens and will be randomly matched with four other participants to form a group of five. Each member of the group is assigned an ID number.
3. At the beginning of each Day, one member of the group will be randomly selected to propose how to split a deduction of (at least) 250 tokens.
4. If three or more members in the group accept the proposal, the proposal is approved, and tokens offered to you will be DEDUCTED from your endowment.
5. If the proposal is rejected, then the group proceeds to the next Meeting of the Day and a proposer will be randomly selected.
6. The volume of the tokens that need to be deducted increases by 20% following each rejection of a proposal in a given Meeting.

Remember that tokens offered to you in the approved proposal are DEDUCTED, not added.

Quiz and Practice Day

To ensure your understanding of the instructions, we will provide you with a quiz below. After the quiz, you will participate in a Practice Day. The Practice Day is part of the instructions and is not relevant to your cash payment. Its objective is to get you familiar with the computer interface and the flow of the decisions in each Meeting. Once the Practice Day is over, the computer will tell you when the official Days begin.

Quiz

To ensure your understanding of the instructions, we ask that you complete a short quiz before we move on to the experiment. This quiz is only intended to check your understanding of the written instructions. It will not affect your earnings. We will discuss the answers after you work on the quiz.

Q1. In each Day, you will be assigned to a group of (A) members. In Meeting 1, each group will decide how to split a deduction of (B) tokens. What are appropriate numbers in (A) and (B)?

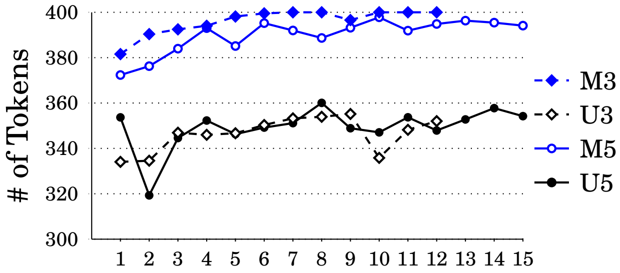
Q2. Suppose that in Day 1, your ID number is 3, and member 1 is selected as a proposer in Meeting 1. Which of the followings is NOT TRUE? (a) If member 1's proposal is rejected, member 1 can be a proposer in Meeting 2. (b) Even if I reject the proposal, it could be approved by majority. (c) In the next Day, my ID number must be 3 again. (d) In Meeting 2 of the current Day, my ID number is unchanged.

Q3. In Meeting 1, there are 250 tokens being divided. Which of the following exemplary proposals makes sense? (a) $(200, 50, 0, 0, 0)$ (b) $(20, 20, 20, 20, 20)$ (c) $(450, -50, -50, -50, -50)$ (d) $(300, 0, 0, 0, 0)$

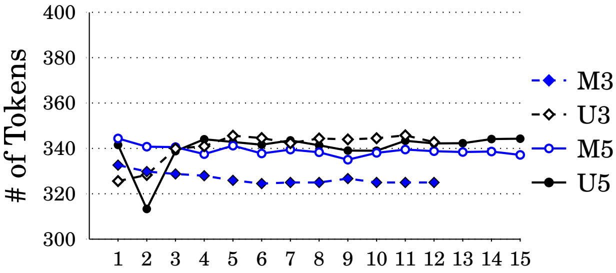
Q4. If a proposal in Meeting 1 is rejected, what will happen next? (a) Your group will move to Meeting 2. One member will be randomly selected as a proposer. (b) Your group will end the Day. The tokens that need to be deducted are equally distributed to each member. (c) The previous proposer will propose one more time. (d) Your group will end the Day. The tokens that need to be deducted will be added to the tokens for the next Day.

Q5. In each Day, you are endowed with 400 tokens. If the approved proposal offered you 100 tokens, what's your earning on that Day?

C Appendix - Supplementary Figures



(a) Proposers



(b) Non-proposers

Figure 7: Average Earnings

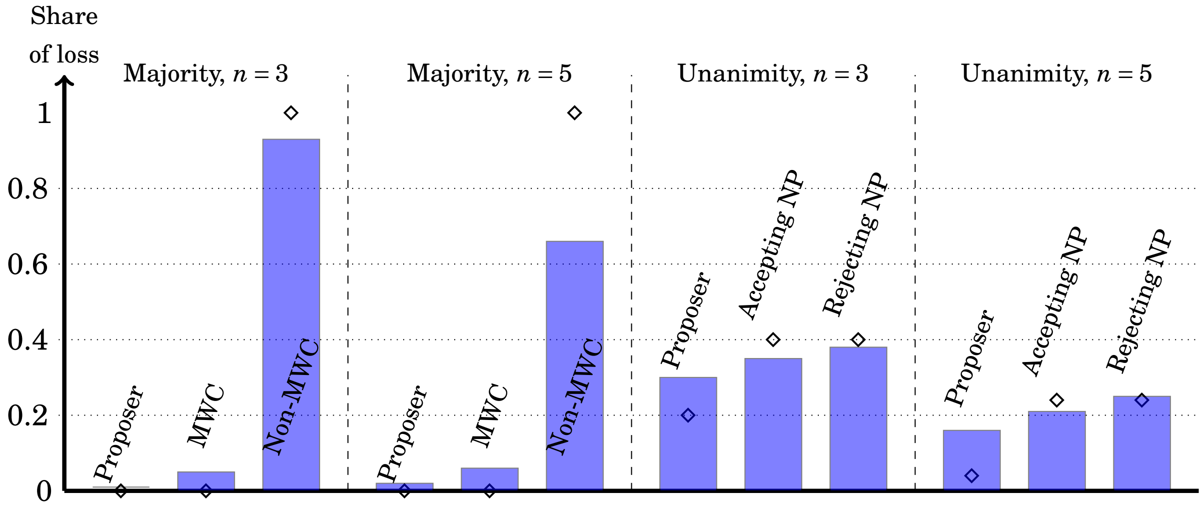
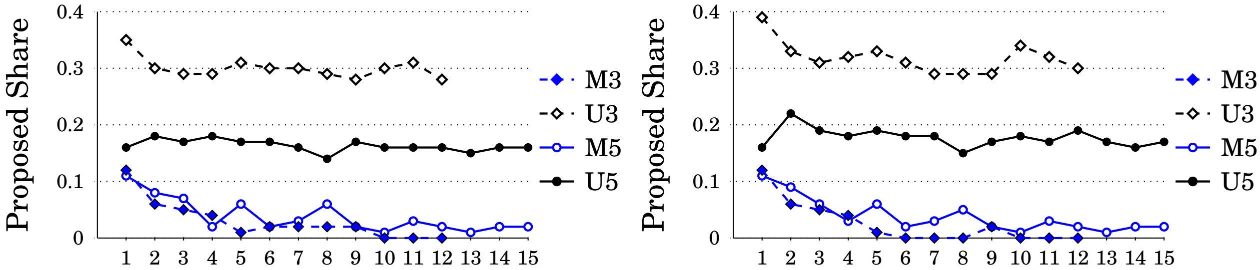


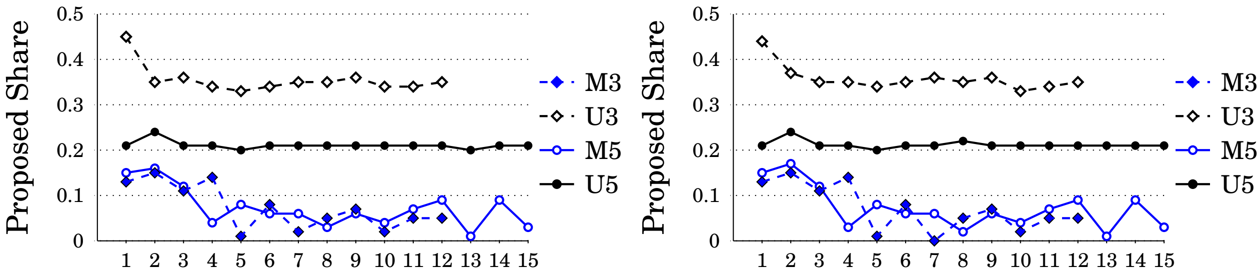
Figure 8: Proposed Shares
All (including rejected) proposals in all periods



(a) All Proposals

(b) Approved Proposals

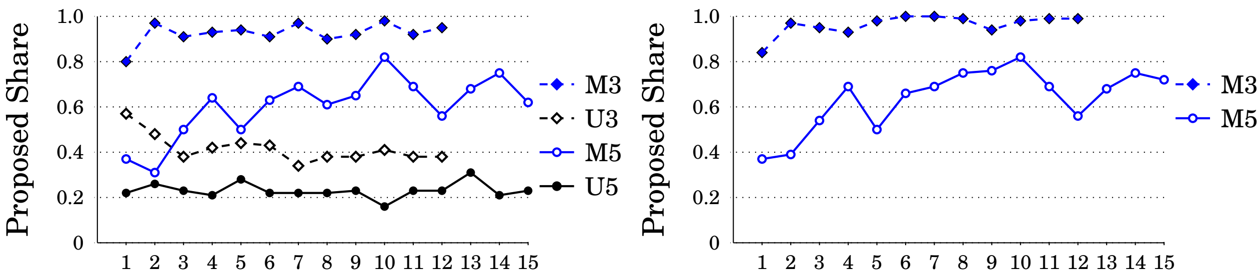
Figure 9: Average Proposed Share - Proposer



(a) All Proposals

(b) Approved Proposals

Figure 10: Average Proposed Share - Accepting Non-proposer



(a) All Proposals

(b) Approved Proposals

Figure 11: Average Proposed Share - Rejecting Non-proposer