On Foundations of Bargaining and Voting

On Foundations of Bargaining and Voting

Habilitation
submitted to the
Department of Management,
Technology and Economics
of ETH Zurich

to obtain the
Venia Legendi in Economics

presented by
Dr. Volker Britz

Zürich, 2019

Acknowledgements

First of all, I would like to thank Prof. Hans Gersbach for his support and advice in the entire process of preparing this Habilitation thesis.

The Habilitation subsumes research on Bargaining and Voting which I have done during my time as a post-doctoral researcher at ETH Zürich and, prior to that, at the Université Catholique de Louvain in Louvain-la-Neuve, Belgium. Each chapter of this thesis was originally written as a scientific paper. I would like to thank my co-authors on these papers, Hans Gersbach, Jean-Jacques Herings, and Arkadi Predtetchinski for fruitful research collaboration. Throughout the years, I have also received many helpful comments and suggestions from the colleagues and academic guests of ETH Zürich and University of Zürich.

Margrit Buser and Christian Waibel have been very generous in their support while I was preparing the Habilitation.

Of course, my deep gratitude goes to my parents, Wolfgang and Dagmar, who have tirelessly extended their encouragement, love, and advice – for my education, my career, and for everything beyond it.

Contents

Introduction, Motivation, and Summary

Why Bargaining Theory ?

One of the fundamental problems in Economics is the following: Suppose that a buyer is willing to pay more for some good or service than it costs a supplier to provide it. In that case, there are "gains from trade." That is, the buyer and seller can divide a surplus among themselves.

How much each party benefits from this surplus, however, depends on the price for which the good or service is exchanged: If that price is close to the buyer's willingness to pay, then the seller gets most of the surplus. If the price is close to the seller's production cost, then the buyer reaps most of the surplus. The key question is: What determines the price, and hence the division of the surplus?

Standard economic theory has a very well-known approach to this question: It describes suppliers' cost of production and buyers' willingness to pay by supply and demand curves. The intersection of these curves determines a market-clearing price. This approach, however, has many limitations. In particular, it relies on the premise that there are sufficiently many essentially identical buyers and suppliers so that none of them can individually influence the price: They are all "price-takers."

In reality, few markets approximate this highly stylized framework: Rather than being price-takers, market participants often do have some degree of market power. While variants of the standard approach can describe situations in which power is concentrated on one side of the market (such as in monopoly or oligopoly), the approach fails when there is some market power on both sides. One classical economic example is the following: A firm may want to hire a worker with some very specific skills, and the worker may want to find a job in a specific town where just one firm demands such specific skills. In that case, neither the worker nor the firm can easily replace one another with some random competitor. Their relationship is, at least to some extent, characterized by a "bilateral monopoly." How are they going to divide the surplus?

Bargaining Theory: The canonical model

Game theorists have developed Bargaining Theory to address this type of surplus division problem between two or more economic agents. The objective is to design a formal game-theoretic model to describe the negotiations that determine surplus division. This approach can be applied whenever two or more parties have access to some surplus, provided they can agree on its division. Bargaining Theory can therefore be used to analyze problems outside the realm of economics. Applications of Bargaining Theory are particularly pertinent in political science. One example is the division of government funds among interests represented by different political parties in a governing coalition. A second example are international negotiations about trade agreements, environmental protection, or use of natural resources.

Much of the game-theoretic literature on bargaining is based on an approach pioneered by Rubinstein (1982): This canonical bargaining model formalizes a negotiation as a sequence of rounds. In each round, one of the players has the exclusive right to make a proposal, which can then be accepted or rejected by the other players. Whenever a proposal is rejected, a new round starts. Opening a new round, however, involves a costly delay. It provides players with an incentive to reach agreement sooner rather than later. In this framework, being the proposer is a source of bargaining power: Loosely speaking, a player who can frequently make proposals, obtains a large share of the surplus.

More formally, the bargaining literature has considered many different "bargaining protocols," that is, rules which specify which player has the right to make the proposal in which round of the canonical bargaining model. Taking a particular bargaining protocol as given, various researchers have investigated the equilibrium division of the surplus.[1] One particularly natural example of a bargaining protocol is the one where a player who rejects the current proposal must make a counter-proposal. Another bargaining protocol often used in the literature is the one where the proposer is drawn from some fixed, time-invariant probability distribution in each round.

Chapter 1 of this Habilitation thesis provides an analysis of an action-dependent protocol, which covers the two aforementioned bargaining protocols as special cases. This chapter therefore provides a substantial generalization of results in the extant literature, and demonstrates how they are related. In particular, Chapter 1 shows how equilibrium

[1] An overview of this literature can be found in Britz et al. (2010).

payoffs converge to the "asymmetric Nash Bargaining Solution." The main idea behind this solution is that it maximizes a weighted product of players' payoffs, where the weights are determined by relative proposal power. In the special case where these weights are equal across players, the Nash Bargaining Solution results. It is a solution concept which was initially introduced in the literature on cooperative games, and goes back to a seminal article by Nobel laureate John Nash, who demonstrated that the Nash Bargaining Solution uniquely satisfies a certain set of desirable properties, see Nash (1953).

Chapter 2 argues, however, that the classes of bargaining protocols studied in Chapter 1 and in Britz et al. (2010) cannot easily be generalized further without invalidating key results in this literature. Thus, Chapters 1 and 2, together with the paper by Britz et al. (2010), can be seen as an exploration of how far some important results in the bargaining literature can be generalized.[2]

Extending the scope of the canonical model

The findings in Chapter 2 underscore the limitations of the canonical approach to bargaining. One important criticism is that the canonical model puts undue emphasis on the precise protocol, that is, on who is allowed to make proposals at what time. While the "right to propose" certainly is an important source of bargaining power, other factors play an important role in actual bargaining situations. One of these factors is a player's ability to credibly commit to certain "red lines." Another important factor is that unfair behavior at the bargaining table may damage trust, and thereby harm the fruitfulness of future cooperation between the players.

In order to fill some of the gaps left by the canonical approach, I have worked extensively on a single-authored line of research. The guiding theme of this research is to extend the scope of bargaining models. To date, this line of research has produced three papers, two of which are part of this Habilitation thesis:

Chapter 3[3] challenges the conventional idea that players find it optimal to make full use of their bargaining power. It provides a formal model in which the players choose how aggressively they wish to bargain. Each individual player can strengthen their bargaining position by being more aggressive, but the size of the surplus is diminished by aggression. – This trade-off is reminiscent of that in social dilemma games. Hence, players have an

[2] Chapter 1 corresponds to the journal article by Britz et al. (2014), while Chapter 2 corresponds to the journal article by Britz et al. (2015).

[3] Chapter 3 corresponds to the journal article by Britz (2018).

incentive to be only "moderately aggressive" at the bargaining table. One interpretation is the following: In actual bargaining situations, overly aggressive behavior by one party may lead the other parties to loose trust, and thereby harm the prospects for cooperation between the parties. As an example, consider again the wage negotiation between a firm and an employee: The surplus that they aim to divide can be thought of as the productivity gain that occurs when this employee works in that firm. If wage bargaining is done in a particularly aggressive way, this may lead to frustration and make the future work relationship less productive: Indeed, the surplus is smaller. The contribution of Chapter 3 is to provide a formal modeling framework for such considerations, and pave the way to incorporating trust between players into a bargaining analysis.

Chapter 4[4] takes issue with the timing of moves in the canonical bargaining model. It is standard to assume that a proposal is made, accepted, and then implemented without any time lapse between these three steps. Only the rejection of a proposal triggers some lapse of time, which is costly to the players due to their impatience. I propose an alternative approach under which a proposal is made, accepted, and implemented at distinct points in time. The motivation is as follows: It may take players an uncertain amount of time to communicate proposals and to "understand" each other's proposals. In practice, the terms of a contract may take time to understand. I explicitly model such a delay, and thereby drive a wedge between the time when a proposal is made and when it is accepted. Moreover, there may be a delay between the time a proposal is accepted and the time it is implemented. In practice, this may be due to institutional constraints: For example, there could be an institutional rule saying that a working contract can only begin on the first day of the month. It would then be irrelevant whether an agreement is reached on, say, January 7th or January 23rd – in either case, it can be implemented only on February 1st. As a result, delay in bargaining may only be costly whenever the end of a month is reached, rather than every time a proposal is rejected. In Chapter 4, I provide an analysis of such bargaining situations, and I discuss the conditions under which the results do or do not resemble those in the canonical bargaining model.

Bargaining and Voting: Open vs. closed rules

Bargaining games as discussed in this Habilitation thesis are closely related to *voting games,* which have the following structure: A player makes a proposal, which is then

[4] Chapter 4 corresponds to the working paper by Britz (2019).

voted on by all players at once. If a certain number of players votes in favor, the proposal is implemented, otherwise, the game proceeds to a new round. The number of affirmative votes required may be a simple majority, or a two-thirds majority, or any other super-majority. If decisions are required to be unanimous, then the voting game reduces to the canonical bargaining game.

The class of bargaining and voting games is often used to model decision-making in parliamentary settings. For instance, each member of parliament wishes to secure subsidies for their own constituency. Such "legislative bargaining" models have become especially popular in the literature starting with the seminal paper of Baron and Ferejohn (1989). They distinguish between *closed rule legislative bargaining* and *open rule legislative bargaining.* The bargaining and voting games discussed so far are examples of the "closed rule": When a proposal is made, it is voted on, and only then can a new proposal be made. In contrast, an "open" rule allows players to make amendments to a proposal before any vote takes place. While open rules are common in real-life parliamentary settings, the bargaining and voting literature has focused primarily on closed rules. The reason is that open rule bargaining is notoriously difficult to study analytically: While the analysis of open rule bargaining in Baron and Ferejohn's original paper is known to be flawed and incomplete, the bargaining literature has so far only made few improvements upon it. The purpose of Chapter 5[5] is to take a new approach to this problem: We define a particular refinement of subgame-perfect equilibrium in stationary strategies, which is, in a precise sense, that equilibrium with the "simplest" possible strategies. This refinement allows us to describe the equilibrium variables of the open rule legislative bargaining game by an analytically tractable equation system. We can establish results that allow for a comparison of the properties of open and closed rule bargaining. One important insight is that equilibrium delays which occur in the open rule bargaining game may be much longer than originally predicted by Baron and Ferejohn (1989) – this suggests that open rules have less desirable properties than previously believed. It is still true, however, that open rule bargaining tends to lead to more egalitarian equilibrium allocations than closed rule bargaining.

[5]Chapter 5 corresponds to the working paper by Britz and Gersbach (2018).

Voting, Information, and Manipulation

Finally, Chapter 6[6] continues the analysis of voting. In particular, this chapter addresses one of the fundamental questions for game-theoretic analyses of the democratic process: *How can democratic procedures and institutions help reveal dispersed information?* Consider a democratic society that has to choose between several alternatives by a voting. Each citizen has some information about the available alternatives. However, no citizen possesses all the information at once, and, moreover, the bits of information available to different citizens may conflict with each other. In an ideal world, all citizens could share their information with each other, and hence make a perfectly informed voting decision. In reality, however, only limited time and resources are available for information sharing. One important challenge in Democracy Research is to design democratic procedures in such a way that the outcome of the procedure is as close as possible to the outcome that would obtain in the hypothetical ideal world with perfect information sharing. The main contribution of Chapter 6 is to study a class of so-called "democratic mechanisms" which guarantees the implementation of a Condorcet winner in a society which chooses from a finite set of alternatives under aggregate uncertainty. The analysis in Chapter 6 is very innovative in so far that it combines voting games and mechanism design, two strands of the Economic Theory literature that are typically seen as separate.

[6]Chapter 6 corresponds to the journal article by Britz and Gersbach (2019).

References

Baron, D. and J. Ferejohn (1989), Bargaining in Legislatures, *American Political Science Review, 83,* 1181-1206.

Britz, V. (2018), Rent-Seeking and Surplus Destruction in Unanimity Bargaining, *Games and Economic Behavior, 109,* 1-20.

Britz, V. (2019), Negotiating with Frictions, CER-ETH Working Paper 19/309.

Britz V. and H. Gersbach (2018), Open Rule Legislative Bargaining, CEPR Discussion Paper DP12966.

Britz, V. and H. Gersbach (2019), Information Sharing in Democratic Mechanisms, *International Journal of Game Theory,* accepted.

Britz, V., P.J.J. Herings, and A. Predtetchinski (2010), Non-cooperative Support for the Asymmetric Nash Bargaining Solution, *Journal of Economic Theory, 145* 1951-1967.

Britz, V., P.J.J. Herings, and A. Predtetchinski (2014), On the Convergence to the Nash Bargaining Solution for Action-Dependent Bargaining Protocols, *Games and Economic Behavior, 86,* 178-183.

Britz, V., P.J.J. Herings, and A. Predtetchinski (2015), Delay, Multiplicity, and Non-existence of Equilibrium in Unanimity Bargaining Games, *Journal of Mathematical Economics, 61,* 192-202.

Nash, J. (1953), Two-Person Cooperative Games, *Econometrica, 21,* 128-140.

Rubinstein, A. (1982), Perfect Equilibrium in a Bargaining Model, *Econometrica, 50,* 97-109.

Chapter 1

On the Convergence to the Nash Bargaining Solution for Action-Dependent Protocols

Joint with Jean-Jacques Herings and Arkadi Predtetchinski
Published in *Games and Economic Behavior*

Abstract: We consider a non–cooperative multilateral bargaining game and study an action–dependent bargaining protocol, that is, the probability with which a player becomes the proposer in a round of bargaining depends on the identity of the player who previously rejected. An important example is the frequently studied rejector–becomes–proposer protocol. We focus on subgame perfect equilibria in stationary strategies which are shown to exist and to be efficient. Equilibrium proposals do not depend on the probability to propose conditional on the rejection by another player. We consider the limit, as the bargaining friction vanishes. In case no player has a positive probability to propose conditional on his rejection, each player receives his utopia payoff conditional on being recognized. Otherwise, equilibrium proposals of all players converge to a weighted Nash Bargaining Solution, where the weights are determined by the probability to propose conditional on one's own rejection.

JEL classification: C78
Keywords: Strategic Bargaining, Subgame Perfect Equilibrium, Stationary Strategies, Nash Bargaining Solution.

Acknowledgements: Financial support of the Netherlands Organization for Scientific Research (NWO) is gratefully acknowleged.

1.1 Introduction

This paper examines the convergence of equilibrium payoffs to the asymmetric Nash bargaining solution in a non–cooperative bargaining game. In contrast to the existing literature on this topic, we allow for the proposer selection process to be action–dependent, that is, influenced by the players' actions throughout the game.

The study of non–cooperative bargaining games has been strongly influenced by Rubinstein (1982). In his seminal paper, the unique division of a surplus among two impatient players is supported by subgame–perfect equilibrium. In unanimity bargaining games with more than two players, Herrero (1985) and Haller (1986) show that the uniqueness of subgame–perfect equilibrium payoffs is not preserved.[1] Therefore, the literature has focused attention on those subgame–perfect equilibria which are in stationary strategies.

The convergence of equilibrium payoffs in the limit as the cost of delay becomes small has been studied for a variety of different bargaining protocols, see Binmore, Rubinstein, and Wolinsky (1986), Hart and Mas-Colell (1996), Laruelle and Valenciano (2008), Miyakawa (2008), Britz, Herings, and Predtetchinski (2010), and Kultti and Vartiainen (2010). All those protocols have in common that they are action–independent: The actions taken by the players in the game have no effect on the identity of the next proposer. We argue that the focus on action–independent protocols alone is a serious limitation because action–dependent protocols are very common in other strands of the bargaining literature. One simple and intuitively appealing example is the protocol where the player who rejects the current proposal is automatically called to make the next proposal. This rejector–becomes–proposer protocol has been introduced in Selten (1981) and has been studied extensively in both the bargaining and the coalition formation literature, see for example Chatterjee, Dutta, Ray, and Sengupta (1993), Bloch (1996), Ray and Vohra (1999), Imai and Salonen (2000), and Bloch and Diamantoudi (2011).

The protocol we study in this paper is more general than the rejector–becomes–proposer protocol. Following Kawamori (2008), we are interested in the case where the identity of the player who rejects a proposal may influence the probability by which a particular player becomes the next proposer. Since now the accept and reject decisions of the players influence the selection of the proposer, this leads to an action–dependent protocol. Such protocols are considerably more difficult to analyze than action–independent ones, and the literature has identified a number of cases where both types of protocol lead to surprisingly different results. For instance, Chatterjee, Dutta, Ray, and Sengupta (1993) provide examples for non-existence of equilibria with immediate agreement in the

[1] One way to restore the uniqueness of subgame–perfect equilibrium is to deviate from unanimous agreement and consider instead a bargaining process where an agreement is reached in several steps and only a subset of the players bargain with each other at each step. Examples of such "partial agreements" can be found in Chae and Yang (1994), Krishna and Serrano (1996), and Suh and Wen (2006). A similar approach has been applied to a coalition formation problem by Moldovanu and Winter (1995).

context of an action–dependent protocol. On the contrary, it has been shown in Okada (1996) that delay cannot occur at equilibrium and in Okada (2011) that equilibria exist when the protocol is action–independent.[2]

Our analysis of stationary subgame perfect equilibria reveals that the set of equilibrium proposals only depends on the bargaining protocol through the probabilities of making counter–offers. A player's probability of making a counter–offer is defined as the probability for that player to become the proposer, given that the previous proposal has been rejected by *this* player.

We show that equilibrium proposals of all players converge to a weighted Nash bargaining solution, where the weights are proportional to the probabilities of making a counter–offer. One exception is the case when the probability of making a counter–offer is zero for all players. In this case the proposer in the initial round obtains his utopia payoff, that is his highest payoff in the set of feasible payoffs that satisfy all the individual rationality constraints. Equilibrium proposals of all players are independent of the continuation probability and do not converge to a common limit.

One implication of our analysis is that the probability of making a counter–offer is a crucial determinant of a player's bargaining power. The existing results on non–cooperative bargaining games are for action–independent protocols only and do not distinguish between the probability of making a proposal and the probability of making a proposal conditional on a rejection. Our paper identifies the latter probabilities as the source of bargaining power.

1.2 The Bargaining Game

We consider a bargaining game between finitely many players in the set $N = \{1, \ldots, n\}$. Each player individually can only attain a disagreement payoff which we normalize to zero. However, the players can jointly achieve any payoff vector v in a set $V \subset \mathbb{R}^n$ if they unanimously agree on such a payoff vector. Each player is assumed to be an expected utility maximizer. The set V of feasible payoffs and the bargaining protocol are the main primitives of the model. We now introduce each in turn.

For vectors u and v in $\mathbb{R}^n$, we write $u \geq v$ if $u_i \geq v_i$ for all $i \in N$, $u > v$ if $u \geq v$ and $u \neq v$, and $u \gg v$ if $u_i > v_i$ for all $i \in N$. A point v of V is said to be *Pareto–efficient* if there is no point u in V such that $u > v$. A point v of V is said to be *weakly Pareto–efficient* if there is no point u in V such that $u \gg v$. We write V_+ to denote the set $V \cap \mathbb{R}^n_+$.

Our first assumption is as follows:

[2]Duggan (2011) presents a very general coalitional bargaining model where equilibrium existence is shown for action–independent protocols. The paper points out that a similar approach to establish equilibrium existence would not work when the protocol is action–dependent.

[A1] The set V is closed, convex, and comprehensive from below. There is a point $v \in V$ such that $v \gg 0$. The set V_+ is bounded. Each weakly Pareto–efficient point of V_+ is Pareto–efficient.

We denote the set of Pareto-efficient points of V by P and write P_+ for the set $P \cap \mathbb{R}^n_+$. A vector $\eta \in \mathbb{R}^n$ is a called a *normal vector* to V at a point $v \in V$ if $(u-v)^\top \eta \leq 0$ for all $u \in V$. In addition, a normal vector to V at v is said to be a *unit normal vector* if $\|\eta\| = 1$.

[A2] There is a continuous function $\eta : P_+ \to \mathbb{R}^n$ such that $\eta(v)$ is a unit normal vector to V at the point v.

Assumption A2 implies that the boundary P_+ does not have kinks. Note that in view of Assumption A1 we have $\eta_i(v) > 0$ for every $i \in N$ such that $v_i > 0$.

Bargaining takes place in discrete time $t = 0, 1, \ldots$. There are $n+1$ probability distributions on the players denoted by $\pi^0, \pi^1, \ldots, \pi^n$, each of which belongs to the unit simplex Δ^n in $\mathbb{R}^n$.

In round $t = 0$, a particular player is chosen as the proposer according to the probability distribution $\pi^0 \in \Delta^n$. The proposer then makes a proposal $v \in V$. Player 1 responds to the proposal by either acceptance or rejection. Once a player $i = 1, \ldots, n-1$ has accepted the proposal, it is the turn of player $i+1$ to accept or to reject.[3] Once player n has accepted the proposal, the game ends and the approved proposal is implemented.

As soon as some player $j \in N$ rejects a proposal in round t, the game ends with probability $1 - \delta > 0$ and payoffs to all players are zero. With the complementary probability δ, the game continues to round $t+1$. The proposer in that round is then drawn from the probability distribution π^j. If the game continues perpetually without agreement, the payoff to every player is zero.

The rejector–becomes–proposer protocol follows from specifying $\pi^i_i = 1$ for all $i \in N$. A polar opposite of the rejector–becomes–proposer protocol, where a rejector proposes with probability zero in the next round, follows by setting $\pi^i_i = 0$ for all $i \in N$. In case $\pi^0, \pi^1, \ldots, \pi^n$ all coincide, we are back in the familiar case of an action–independent protocol with time–invariant recognition probabilities.

It is well–known that bargaining games with more than two players admit a wide multiplicity of subgame–perfect equilibria (SPE), see Herrero (1985) and Haller (1986). We will restrict attention to subgame–perfect equilibria in stationary strategies (SSPE). A *stationary strategy* for player i consists of a proposal $\theta^i \in V$ which i makes whenever it is his turn to propose and an acceptance set $A^i \subset V$ which consists of all the proposals which player i would be willing to accept if they were offered to him. We denote the *social acceptance set* by $A = \cap_{i \in N} A^i$ and write the profile of stationary strategies $(\theta^1, A^1, \ldots, \theta^n, A^n)$ more concisely as $(\Theta, \mathcal{A})$.

[3] Throughout the paper, for the sake of simplicity, we assume that players respond to a proposal in the fixed order $1, \ldots, n$. All results would carry over to the case with arbitrary voting orders.

Consider some stationary strategy profile $(\Theta, \mathcal{A})$. The vector of *continuation payoffs* after i's rejection, denoted q^i, is the vector of payoffs in the subgame that follows the rejection of a proposal by some player $i \in N$. It is well–defined since $(\Theta, \mathcal{A})$ is stationary. Crucial for our analysis is the vector $r = (q_1^1, \ldots, q_n^n)$ of *reservation payoffs*. Under a protocol with time–invariant recognition probabilities, i.e. when $\pi^1 = \cdots = \pi^n$, one and the same vector of continuation payoffs would result no matter which player rejected the current proposal, i.e. $q^1 = \cdots = q^n$. Consequently, the reservation and continuation payoff vectors are equal to each other. If, however, we allow for an action–dependent protocol, the reservation payoff vector is in general not equal to any of the continuation payoff vectors. This disparity between the reservation and continuation payoffs is what makes the analysis of action–dependent protocols different from the analysis of action–independent protocols.

1.3 Convergence of SSPE Payoffs

In order to study the convergence of SSPE payoffs when the breakdown probability tends to zero, we first state an existence result for SSPEs and a characterization of the associated proposals.

Theorem 1.1. *An SSPE exists. In every SSPE, agreement is reached immediately, and the proposals are such that for every $i \in N$,*

$$\theta^i \in P_+, \tag{1.1}$$

$$\theta_j^i = \alpha_j \theta_j^j, \quad j \in N \setminus \{i\}, \tag{1.2}$$

where $\alpha_j = \delta\pi_j^j/(1-\delta+\delta\pi_j^j)$. Conversely, for every profile of proposals $(\theta^1, \ldots, \theta^n)$ satisfying (1.1)–(1.2), there is an SSPE with this profile of proposals.

Proof: For a detailed derivation of the existence and characterization theorem, we refer to the working paper by Britz, Herings, and Predtetchinski (2013). For existence of an SSPE, one has to show that the system consisting of equations (1.1) − (1.2) has a solution. The argument can be easily adapted from Banks and Duggan (2000). The proof that for every profile of proposals $(\theta^1, \ldots, \theta^n)$ satisfying (1.1)–(1.2) there is an SSPE with this profile of proposals is standard but tedious. Here, we restrict ourselves to the proof of the immediate acceptance property and the derivation of (1.1)–(1.2), where we focus on the parts of the proof which are different from the action-independent case. In what follows we fix an SSPE $(\Theta, \mathcal{A})$.

Step 1: *For each $v \in V$: If $v \in V$ is such that $v \gg r$, then $v \in A$. If $v \in A$, then $v \geq r$.*

The proof of this step is easy and therefore omitted.

Step 2: *The social acceptance set A is non–empty and $r \geq 0$. There exists $v \in V$ such that $v \gg r$.*

Suppose there is no $v \in V$ such that $v \gg r$. In view of Assumption A1, there is no $v \in V$ such that $v > r$. It now follows from Step 1 that $A \subset \{r\}$. First suppose that $A = \emptyset$. In this case equilibrium strategies lead to payoffs of zero for all players, so $r = 0$. Under Assumption A1 there is a vector $v \in V$ with $v \gg 0$, a contradiction to our supposition. Hence $A = \{r\}$.

Then, after a rejection, only two outcomes are possible: Either agreement on r is reached at some future time or zero payoffs result. The vector of players' continuation payoffs after a rejection is therefore a convex combination of 0 and r, where the former has a weight of at least $1 - \delta$. But this implies $r_i \leq \delta r_i$ for all $i \in N$. Since $\delta < 1$, we conclude that $r = 0$. As before, this leads to a contradiction.

We conclude that there is $v \in V$ such that $v \gg r$. Step 1 implies that $v \in A$. Since each player can choose to reject all proposals, $r \geq 0$.

STEP 3: *For every $i \in N$, $\theta^i \in A$.*

Let u_i be the SSPE utility to player i at a history where it is player i's turn to make a proposal. It holds that $u_i = \theta^i_i$ if $\theta^i \in A$ and $u_i = q^j_i$ if $\theta^i \notin A$, where j is the least element of N such that $\theta^i \notin A^j$.

By making a proposal $v \in A$, player i guarantees himself a payoff of v_i. It follows that $u_i \geq v_i$ for every $v \in A$. In particular $u_i > 0$ since by Step 2 there is a vector $v \in A$ such that $v_i > 0$.

Let $U = \{0\} \cup (A \cap \{\theta^1, \ldots, \theta^n\})$. This is the set of all possible outcomes of the game if play follows the strategy $(\Theta, \mathcal{A})$. We know from the preceding paragraph that $u_i \geq v_i$ for all $v \in U$ and that $u_i > 0$. Take any $k \in N$. The vector q^k of continuation payoffs is a convex combination of the vectors in U, with 0 having a weight of at least $1 - \delta$. It follows that $u_i > q^k_i$. Since this holds for each $k \in N$, we must have $u_i = \theta^i_i$ and $\theta^i \in A$.

STEP 4: *For every $i \in N$, $\theta^i \in P_+$ and $\theta^i_j = r_j$ for every $j \in N \backslash \{i\}$.*

The proof follows a standard argument and is therefore omitted.

STEP 5: *For every $i \in N$, for every $j \in N \setminus \{i\}$, $\theta^i_j = \alpha_j \theta^j_j$.*

Since each proposal belongs to the social acceptance set, the reservation payoffs can be computed as follows:

$$r_j = \delta \sum_{k=1}^{n} \pi^j_k \theta^k_j = \delta \pi^j_j \theta^j_j + \delta(1 - \pi^j_j) r_j.$$

Solving for r_j, we see that $r_j = \alpha_j \theta^j_j$. Combining this with Step 4 we obtain the result. □

In the existing literature on n–player bargaining games, it has been shown for a variety of action–independent protocols that the SSPE proposals of all players converge to a common limit proposal. We are interested in the question whether this result carries over to an action–dependent protocol. Indeed, it turns out that the existence of a limit equilibrium proposal is preserved under an action–dependent protocol, except in the degenerate case where the probability of proposing after one's own rejection is zero for all players.

For $m \in \mathbb{N}$, let $(\delta_m)_{m\in\mathbb{N}}$ be a sequence of continuation probabilities converging to 1 and let Θ_m be SSPE equilibrium proposals of the game with continuation probability δ_m. Equilibrium proposals are said to converge to each other if for all $i, j \in N$ it holds that $\|\theta_m^i - \theta_m^j\|_\infty \to 0$.

Proposition 1.1.

1. *If* $\pi_j^j = 0$ *for all* $j \in N$, *then equilibrium proposals do not converge to each other.*
2. *If there is* $j \in N$ *such that* $\pi_j^j > 0$, *then equilibrium proposals do converge to each other.*

Proposition 1.1 can be derived from the SSPE characterization. We define $\hat{v}$ as the vector of *utopia payoffs*, where the utopia payoff of a player $i \in N$ is the highest payoff in V for player i that satisfies all the individual rationality constraints, that is $\hat{v}_i = \max\{v_i \in \mathbb{R} \mid v \in V_+\}$. Consider the case where $\pi_j^j = 0$ for all $j \in N$. In that case, we have $\alpha_j = 0$ and hence $r_j = 0$ for all $j \in N$. It follows that $\theta_j^i = 0$ for all $j \in N$ and $i \in N \setminus \{j\}$. Due to the Pareto–efficiency of all proposals, it must be true that $\theta_j^j = \hat{v}_j$ for all $j \in N$. This argument does not depend on the value of δ. SSPE proposals do not converge to each other. Now consider the case where there exists $j \in N$ such that $\pi_j^j > 0$. Then, we see that α_j converges to one in the limit as δ goes to one. Thus in the limit we have that all players $j \in N$ such that $\pi_j^j > 0$ are offered the same payoff by every proposer. All other players are offered a zero payoff when they are responding, and the Pareto–efficiency of all proposals implies that they also receive zero when they are proposers.

Every accumulation point of SSPE proposals when δ tends to one is called a *limit equilibrium proposal.* We proceed by showing that if there is $j \in N$ such that $\pi_j^j > 0$, then the limit equilibrium proposal is unique and equal to the asymmetric Nash bargaining solution where player i has weight π_i^i. Given a vector $\lambda \in \mathbb{R}_+^n \setminus \{0\}$, we define the λ–Nash product $\rho_\lambda : \mathbb{R}_+^n \to \mathbb{R}$ by

$$\rho_\lambda(v) = \prod_{i\in N} v_i^{\lambda_i}, \quad v \in \mathbb{R}^n.$$

Definition 1.1. *Given a vector* $\lambda \in \mathbb{R}_+^n \setminus \{0\}$, *the maximizer of the function* ρ_λ *on* V_+ *is called the* λ*–Nash bargaining solution.*

Under our assumptions, the maximizer of the function ρ_λ on V_+ is indeed unique. It is a Pareto–efficient point of V which is uniquely characterized by the following conditions: for each i and j in N,

$$\text{if } \lambda_i = 0, \text{ then } v_i = 0, \tag{1.3}$$

$$\text{if } \lambda_i, \lambda_j > 0, \text{ then } \frac{v_i \eta_i(v)}{\lambda_i} = \frac{v_j \eta_j(v)}{\lambda_j}. \tag{1.4}$$

Theorem 1.2.

1. *If* $\pi_j^j = 0$ *for all* $j \in N$, *then each player's expected payoff is* $\pi_j^0 \hat{v}_j$.

2. *If there is* $j \in N$ *such that* $\pi_j^j > 0$, *then the limit equilibrium proposal is unique and is equal to the* $(\pi_1^1, \ldots, \pi_n^n)$*–Nash bargaining solution.*

Proof: The first part follows from the Proposition 1.1, we show the second part. We verify that each limit equilibrium proposal satisfies the conditions (1.3)–(1.4) with $\lambda_i = \pi_i^i$. Let

$$\tilde{N} = \{i \in N \mid \pi_i^i > 0\}.$$

Let $(\theta^1, \ldots, \theta^n)$ be SSPE proposals in a game with continuation probability δ. By the definition of the normal vector it holds for any two players i and j that

$$(\theta^j - \theta^i)^\top \eta(\theta^i) \leq 0.$$

Note that the proposals θ^i and θ^j can only differ in components i and j. Solving for the inner product, we can therefore rewrite the previous inequality as

$$(\theta_j^j - \theta_j^i)\eta_j(\theta^i) + (\theta_i^j - \theta_i^i)\eta_i(\theta^i) \leq 0.$$

Substituting for θ_j^i and θ_j^j from equation (1.2) and dividing by $1-\delta$ yields

$$\frac{\theta_j^j \eta_j(\theta^i)}{1-\delta+\delta\pi_j^j} \leq \frac{\theta_i^i \eta_i(\theta^i)}{1-\delta+\delta\pi_i^i}.$$

Let $\bar{\theta}$ be a limit equilibrium proposal. Taking the limit of the latter inequality along a sequence of equilibrium proposals converging to $\bar{\theta}$, we obtain for all $i, j \in \tilde{N}$,

$$\frac{\bar{\theta}_j \eta_j(\bar{\theta})}{\pi_j^j} \leq \frac{\bar{\theta}_i \eta_i(\bar{\theta})}{\pi_i^i}.$$

Interchanging the roles of the players i and j, we obtain the equality

$$\frac{\bar{\theta}_j \eta_j(\bar{\theta})}{\pi_j^j} = \frac{\bar{\theta}_i \eta_i(\bar{\theta})}{\pi_i^i}, \quad i, j \in \tilde{N}. \tag{1.5}$$

This shows that $\bar{\theta}$ satisfies (1.4). The fact that $\bar{\theta}$ satisfies (1.3) follows at once from the fact that a player $k \in N$ with $\pi_k^k = 0$ receives a zero payoff irrespective of δ so that $\bar{\theta}_k = 0$. □

It is remarkable that the probability to become the proposer after another player's rejection does not matter for the limit equilibrium payoffs. This is in contrast to the analogous results for action–independent protocols. For instance, Britz, Herings, and Predtetchinski (2010) study a multilateral bargaining game in which the proposer is selected according to a Markov process. That is, the probability distribution from which the proposer at round $t+1$ is drawn depends on the identity of the proposer in round t. In that case, the limit equilibrium corresponds to an asymmetric Nash bargaining solution where the vector of bargaining weights is given by the stationary distribution of the Markov process. This Markov process is represented by an $n \times n$–matrix, and all entries of this matrix influence the limit equilibrium payoffs.

1.4 Conclusion

We have considered multilateral bargaining games with action–dependent protocols. The identity of the player who rejects the current proposal determines the probability distribution from which the next proposer is drawn. The probability with which a player proposes after his own rejection is crucial for the equilibrium prediction. In particular, if all players have the same positive probability of proposing after their own rejection, we find convergence to the (symmetric) Nash bargaining solution.

One rather surprising outcome of our analysis is that the probability to become proposer conditional on the rejection of another player does not affect equilibrium payoffs at all. This also sheds new light on the findings in Miyakawa (2008) and Laruelle and Valenciano (2008) concerning the protocol with time–invariant recognition probabilities, which is a special case of our model. Under time–invariant recognition probabilities, it is impossible to discern the effect of the probability to become proposer after one's own rejection as opposed to the probability to propose after another player's rejection. Our more general setup makes the importance of this distinction apparent.

We show that the convergence to the asymmetric Nash bargaining solution breaks down in the degenerate case where each player has zero probability of proposing after his own rejection. In that case, the equilibrium payoffs are determined by the utopia point and the recognition probabilities in the initial round.

References

Banks, J., and J. Duggan (2000), "A Bargaining Model of Collective Choice," *American Political Science Review, 94,* 73–88.

Binmore, K., A. Rubinstein, and A. Wolinsky (1986), "The Nash Bargaining Solution in Economic Modelling," *Rand Journal of Economics, 17,* 176–188.

Bloch, F. (1996), "Sequential Formation of Coalitions in Games with Externalities and Fixed Payoff Division," *Games and Economic Behavior, 14,* 90–123.

Bloch, F., and E. Diamantoudi (2011), "Non-cooperative Formation of Coalitions in Hedonic Games," *International Journal of Game Theory, 40,* 263–280.

Britz, V., P.J.J. Herings, and A. Predtetchinski (2010), "Non-cooperative Support for the Asymmetric Nash Bargaining Solution," *Journal of Economic Theory, 145,* 1951–1967.

Britz, V., P.J.J.Herings, and A.Predtetchinski (2013), "On the Convergence to the Nash Bargaining Solution for Action–dependent Protocols," CORE Discussion Paper 2013/44.

Chae, S., and J.-A. Yang (1994), "An N-Person Pure Bargaining Game," *Journal of Economic Theory, 62,* 86–102.

Chatterjee, K., B. Dutta, D. Ray, and K. Sengupta (1993), "A Non-cooperative Theory of Coalitional Bargaining," *Review of Economic Studies, 60,* 463–477.

Duggan, J. (2011), "Coalitional Bargaining Equilibria," Working Paper, 1–33.

Haller, H. (1986), "Non-cooperative Bargaining of $N \geq 3$ Players," *Economics Letters, 22,* 11–13.

Hart, S., and A. Mas-Colell (1996), "Bargaining and Value," *Econometrica, 64,* 357–380.

Herrero, M.J. (1985), *A Strategic Bargaining Approach to Market Institutions,* PhD Thesis, London School of Economics.

Imai, H., and H. Salonen (2000), "The Representative Nash Solution for Two-sided Bargaining Problems," *Mathematical Social Sciences, 39,* 349–365.

Kawamori, T. (2008), "A Note on Selection of Proposers in Coalitional Bargaining," *International Journal of Game Theory, 37,* 525–532.

Krishna, V., and R. Serrano (1996), "Multilateral Bargaining," *Review of Economic Studies, 63,* 61–80.

Kultti, K., and H. Vartiainen (2010), "Multilateral Non-Cooperative Bargaining in a General Utility Space," *International Journal of Game Theory, 39,* 677–689.

Laruelle, A., and F. Valenciano (2008), "Noncooperative Foundations of Bargaining Power in Committees and the Shapley-Shubik Index," *Games and Economic Behavior, 63,* 341–353.

Miyakawa, T. (2008), "Noncooperative Foundation of n-Person Asymmetric Nash Bargaining Solution," *Journal of Economics of Kwansei Gakuin University, 62,* 1–18.

Moldovanu, B. and E. Winter (1995), "Order–independent equilibria," *Games and Economic Behavior, 9,* 21–34.

Okada, A. (1996), "A Noncooperative Coalitional Bargaining Game with Random Proposers," *Games and Economic Behavior, 16,* 97–108.

Okada, A. (2011), "Coalitional Bargaining Games with Random Proposers: Theory and Application," *Games and Economic Behavior, 73,* 227–235.

Ray, D., and R. Vohra (1999), "A Theory of Endogenous Coalition Structures," *Games and Economic Behavior, 26,* 286–336.

Rubinstein, A. (1982), "Perfect Equilibrium in a Bargaining Model," *Econometrica, 50,* 97–109.

Selten, R. (1981), "A Noncooperative Model of Characteristic Function Bargaining," in V. Böhm and H.H. Nachtkamp (eds.), *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern,* Bibliografisches Institut Mannheim, pp. 131–151.

Suh, S.-C., and Q. Wen (2006), "Multi-agent Bilateral Bargaining and the Nash Bargaining Solution," *Journal of Mathematical Economics, 42,* 61–73.

Chapter 2

Delay, Multiplicity, and Non-Existence of Equilibrium in Unanimity Bargaining Games

Joint with Jean-Jacques Herings and Arkadi Predtetchinski
Published in the *Journal of Mathematical Economics*

Abstract: We consider a class of perfect information bargaining games with unanimity acceptance rule. The proposer and the order of responding players are determined by the state that evolves stochastically over time. The probability distribution of the state in the next period is determined jointly by the current state and the identity of the player who rejected the current proposal. This protocol encompasses a vast number of special cases studied in the literature. We show that subgame perfect equilibria in pure stationary strategies need not exist. When such equilibria do exist, they may exhibit delay. Limit equilibria as the players become infinitely patient need not be unique.

JEL classification: C72
Keywords: Strategic Bargaining, Subgame perfect equilibrium, Stationary Strategies, Nash bargaining solution

2.1 Introduction

In his seminal paper, Rubinstein (1982) studies the division of a surplus among two impatient players through a non-cooperative bargaining game. Following this contribution, a rich literature has emerged which extends and generalizes Rubinstein's approach. A number of general results have persistently and recurrently emerged from this literature: in bargaining games where a surplus is divided under unanimity rule, equilibria exist, are efficient, and converge to a weighted Nash bargaining solution in the limit. In this paper, we explore the boundaries of the scope under which these general results are valid.

We consider bargaining games with the following characteristics. There is a finite number of players who need to make a unanimous choice for one particular payoff vector within a full-dimensional set of feasible payoffs. The game is set in discrete time and in each round of the game, one player is selected as the proposer. His role is to suggest one particular feasible payoff vector. The other players then sequentially accept or reject this proposal in some fixed order. If all players agree to the proposal, the game ends and the agreed upon payoffs are realized. As soon as one of the players rejects the current proposal, the game proceeds to the next round. We assume an exogenous breakdown of the negotiation to occur after each disagreement with probability $1-\delta$. Time-discounting and the possibility of an exogenous breakdown are largely interchangeable interpretations of δ. The term "bargaining friction" can be used to capture both of them. The importance of the bargaining friction lies in the fact that it creates an incentive to come to an agreement sooner rather than later.

In order to complete the description of a unanimity bargaining game, one has to specify a rule which determines which player is the proposer in what round. We will refer to this rule as the protocol in the sequel. Rubinstein (1982) studies a game with only two players who simply take turns in making proposals, the alternating offer protocol. Rubinstein finds a unique subgame perfect equilibrium. The equilibrium strategies happen to be stationary. It is well-known that the uniqueness of subgame perfect equilibrium breaks down in unanimity bargaining games with more than two players. With regard to those games, the literature focuses on subgame perfect equilibrium in pure stationary strategies (SSPE), which allow sharp predictions of the equilibrium payoffs at least when the discount factor is sufficiently close to one. Arguably the most obvious generalization of Rubinstein's alternating offer protocol to the case with more than two players is the rotating protocol, under which players become proposers in ascending order, and the first player proposes again after the last player. One alternative proper generalization of the alternating offers protocol is the rejector-proposes protocol introduced in Selten (1981) in a coalitional bargaining set-up. Under that rule, the first player to reject the current proposal becomes the next proposer. The rejector-proposes protocol is an example of an endogenous protocol in which the actions taken by the players throughout the game have an influence on the proposer selection. One important example of a protocol which is not a proper generalization of Rubinstein's alternating offers protocol is the time-invariant probability protocol which consists of an exogenously given probability distribution from which the proposer is drawn in each round.

The literature on unanimity bargaining games has established some results that are generally valid no matter which of these protocols is assumed.

1. An SSPE exists.

2. Every SSPE has no delay.

3. Every SSPE has efficient proposals.

4. There is a unique limit equilibrium as δ approaches one.

5. All limit equilibrium proposals are equal to a weighted Nash bargaining solution.

The aforementioned results follow from Britz, Herings, and Predtetchinski (2010, 2014) for very general bargaining protocols, and have been shown for some more specific protocols by Binmore, Rubinstein, and Wolinsky (1986), Hart and Mas-Colell (1996), Kultti and Vartiainen (2010), Laruelle and Valencioano (2008), and Miyakawa (2008). The weights in the Nash bargaining solution corresponding to the limit equilibrium proposals depend on the distribution of bargaining power inherent in the protocol. Kultti and Vartiainen (2010) show that the weights are all equal to each other under the rotating protocol. Miyakawa (2008) and Laruelle and Valenciano (2008) study the time-invariant probability protocol. In this case, the vector of bargaining weights is given by the time-invariant probability distribution. Britz, Herings, and Predtetchinski (2010) study a protocol where the selection of the proposer is described by a Markov process. That is, there are n probability distributions on the n players. The identity of the proposer in the current round determines which of the probability distributions is used to draw the proposer in the following round. This Markovian protocol is both a generalization of the time-invariant probability protocol and the rotating protocol. Then, the vector of bargaining weights is given by the stationary distribution of the Markov process.

Britz, Herings, and Predtetchinski (2014) study a class of endogenous protocols, thereby covering the rejector-proposes protocol. They consider protocols which consist of n probability distributions on the n players. The identity of the player who rejects the current proposal determines which of those probability distributions will be used to draw the following proposer. The vector of bargaining weights is shown to be proportional to the vector of probabilities with which the players propose after their own rejections.

In this paper, we present a class of unanimity bargaining games that allows for a rich family of bargaining protocols. This is achieved by introducing a finite set of states and for each state, conditional on the identity of the rejecting player, a vector of transition probabilities to the new states. The state determines the identity of the proposer and the order of the responses. It is easily verified that all the aforementioned protocols follow as special cases. The modeling approach is closely related to the one of Merlo and Wilson (1995). We are more general in allowing for endogenous protocols, i.e. the vector of transition probabilities may depend on the identity of the rejecting player. Motivated by our desire to study the effects of the bargaining protocol itself, we are less general in not allowing for the set of feasible payoffs to depend on the state.

We demonstrate that the results of the bargaining literature as enumerated above do not generalize, even when we require the set of feasible payoffs to correspond to the division of a unit surplus. We construct an example with three players and three states. A player is the proposer in his own state and the states corresponding to Players 2 and 3 are absorbing. Once Player 2 or Player 3 is selected as the proposer, he will remain the proposer forever. In state 1, the protocol follows the rejector-proposes protocol after rejections by Players 2 or 3. In the example, any SSPE predicts delay. SSPE proposals need not be efficient. In the example there is a continuum of SSPEs. Since the example is valid for arbitrarily high values of δ, it is then used to show that limit equilibria are not unique, that there are two accumulation points of limit equilibrium utilities, and that limit equilibrium proposals may not be equal to each other. We show that the example is robust to perturbations of the transition probabilities.

The main intuition behind the example is that Players 2 and 3 capture the entire surplus in their own state. Since the protocol is of the rejector-proposes type in state 1, an SSPE with immediate acceptance requires Player 1 to offer both of them at least δ. Since the total surplus is equal to one, this is clearly infeasible when δ is above one half. All SSPE are therefore such that the offer by Player 1 in state 1 is rejected by one of the other players.

Yildiz (2003) studies the role of optimism in explaining bargaining delay. He has an example which is similar to ours in the sense that each of the responding players has a continuation payoff of δ, so that immediate agreement is impossible when δ exceeds one half. In the example by Yildiz (2003), however, this result is driven by the fact that due to optimism each player believes that he will become the proposer in the next round with probability one. These beliefs are incompatible and therefore at least one player is wrong about the protocol. In contrast, the delay in our example is derived in a set-up where the bargaining protocol is common knowledge among the players.

Finally, we modify our leading example in a simple way. In state 1, we use the rejector-proposes protocol for Players 2 and 3 with probability one half and assume that Player 1 remains the proposer with the complementary probability. We demonstrate that now both SSPEs with immediate agreement and SSPEs with delay fail to exist. Herings and Predtetchinski (2009) have shown for the time-invariant probability protocol that SSPEs exist even when the set of feasible payoffs is non-convex. It follows that variations in the protocol are more problematic for fundamental properties like existence and efficiency of SSPEs than variations in the set of feasible payoffs.

Our results complement some of the examples of equilibrium delay and non-existence found in the literature. An example of an SSPE exhibiting delay has been given in Chatterjee, Dutta, Ray, and Sengupta (1993) in the context of coalitional bargaining. Unlike the unanimity bargaining games considered here, in coalitional bargaining games a proposing player may choose to make an offer to a subset of the players. The approval of the proposal by all players in the chosen coalition is then sufficient for the proposal to pass. Also in a coalitional bargaining context, Bloch (1996) shows that SSPEs need not exist. Merlo and Wilson (1995) show that an SSPE may exhibit delay if the size of the cake changes stochastically over time. Jéhiel and Moldovanu (1995) show that delay can arise due to externalities. In addition to these examples, where delay arises in a complete and perfect information framework, there is a literature on bargaining delays when the parties are asymmetrically informed, see for instance the review by Ausubel, Cramton, and Deneckere (2002).

The plan of the paper is as follows. We start by formally describing a class of unanimity bargaining games in Section 2. Section 3 summarizes the results in the literature regarding existence, immediate agreement, efficiency, and limit equilibria. Section 4 presents the example where an SSPE predicts delay, inefficiency, and non-uniqueness of the limit equilibrium. Section 5 shows the example to be robust to perturbations in the transition probabilities. We show in Section 6 that an SSPE may even fail to exist all together. Section 7 concludes.

2.2 Model

We consider a non-cooperative bargaining game $G(N, V, S, \iota, p^0, p, \delta)$, where $N = \{1, \ldots, n\}$ is the set of players and $V \subset \mathbb{R}^n$ is the set of feasible payoffs. Bargaining takes place in discrete time $t = 0, 1, \ldots$. In each round, one player is selected as the proposer and proposes an element v of V. Next, the players sequentially respond to the proposal and, in case of unanimous acceptance, the proposal is implemented and the game ends with payoffs v to the players. As soon as one player rejects, the game breaks down with probability $1 - \delta$, and continues to the next round with probability $\delta \in (0, 1)$. In case of breakdown, as well as in case of perpetual disagreement, payoffs to all players are equal to zero.

Our emphasis will be on the role of the protocol in determining the bargaining outcome. The set of feasible payoffs V is therefore kept fixed in each round, but the bargaining protocol is allowed to be quite general. To achieve this, we make use of a finite state space S. The function $\iota : S \to N \times \Pi$, where Π is the set of permutations on N, assigns to each state a proposer and an order of responders. That is, if $\iota(s) = (i, \pi)$, then Player i is the proposer in state s and all players sequentially respond to the proposal in the order

$\pi(1), \ldots, \pi(n)$ given by the permutation π.

In round $t = 0$, the initial state is determined by the probability distribution $p^0 \in \Delta(S)$, where $\Delta(S)$ is the set of probability distributions on S. In any round $t > 0$, the state of the game is determined by *transition functions* $p^j : S \to \Delta(S)$, one for each Player $j \in N$. If Player $j \in N$ rejects the proposal at time t when the game is in state s, then $p^j(s)$ returns the probability distribution from which the state at time $t+1$ is drawn conditional on the continuation of the negotiations.

Many protocols that have been studied in the bargaining literature are special cases of the class of protocols described above, up to relatively unimportant modeling details. Such modeling details concern whether there is some probability of breakdown of negotiations, or whether players have time preferences. In case time preferences take the discounted utility form, although conceptually different from the risk preferences that are needed to study models with breakdown, both approaches lead to the same results as argued in Binmore, Rubinstein, and Wolinsky (1986). Another issue is whether players vote simultaneously or sequentially. Under simultaneous voting, the solution concept of subgame perfection has less bite, and on top of subgame perfection it is typically required that players use stage-undominated voting strategies to avoid coordination problems. We study sequential voting in this paper. A final modeling detail is that we formally allow the proposer to turn down his own proposal.

The alternating offer protocol studied in the seminal contribution of Rubinstein (1982) corresponds to the case where $S = N = \{1, 2\}$, Player s is the proposer in state s, and state transitions are such that the state alternates between periods. Player 1 is the initial proposer. Kultti and Vartiainen (2010) consider a multilateral extension of alternating offer bargaining, where proposers rotate in making offers. Their model corresponds to the case where $S = N$, Player s is the proposer in state s, and the state transition is to state $s + 1$ modulo n with probability 1 if the current state is s. Binmore (1987) and Banks and Duggan (2000) consider the time-invariant probability protocol. This protocol results when $S = N$ and there is a fixed probability distribution p^0 on S such that the proposer is selected in accordance with p^0 in every time period. Kalandrakis (2004) and Britz, Herings, and Predtetchinski (2010) consider the case where $S = N$ and require that for all $j, k \in N$, $p^j = p^k$. For $s \in S$, it holds that $\iota(s) = (s, \pi^0)$, where π^0 is the identity. The state denotes the current proposer and Player i responds before Player j if and only if $i < j$. State transitions are not influenced by the identity of the rejecting player, but are otherwise general, so this Markovian protocol includes the rotating protocol and the time-invariant probability protocol as special cases. Merlo and Wilson (1995) consider a general state space S and assume that for all $j, k \in N$, $p^j = p^k$. They allow the set of feasible payoffs to depend on the state s, but since our attention here is on the influence of the protocol on the allocation of payoffs, we consider a fixed set V instead.

All the bargaining protocols described in the previous paragraph have in common that the actions taken by the players are without consequence for the way the bargaining protocol proceeds in case of a rejection, i.e. for all $j, k \in N$ it holds that $p^j = p^k$. We refer to these protocols as exogenous.

The rejector-proposes protocol is introduced in Selten (1981) in a coalitional bargaining set-up and specifies that the player who rejects the current proposal is automatically called upon to make the next proposal. Kawamori (2008) generalizes this protocol to allow for a general probabilistic selection of a new proposer, conditional on who rejects the current proposal. When we apply his coalitional bargaining model to our unanimity bargaining set-up, we obtain the case where $S = N$, for $s \in S$ it holds that $\iota(s) = (s, \pi^0)$, where π^0 is some fixed permutation of the players, and for all $s, s' \in S$, $p^j(s) = p^j(s')$.

Whenever for some $j, k \in N$, $p^j \neq p^k$, the actions of the players influence the way the bargaining protocol proceeds, and we refer to such protocols as endogenous protocols. The rejector-proposes protocol is a key example of an endogenous protocol.

2.3 Results in the Existing Literature

Multilateral bargaining games are known to admit a wide multiplicity of subgame perfect equilibria, see Herrero (1985) and Haller (1986). It is therefore common in the literature to restrict attention to subgame perfect equilibria in pure stationary strategies. Although ideally the notion of stationarity should follow endogenously from the specification of the game as in Maskin and Tirole (2001), the literature typically takes the more ad hoc approach described below, which in general is weaker than the stationarity notion of Maskin and Tirole (2001).

A *pure stationary strategy* for Player i in the game $G(\delta)$ consists, for each state $s \in S$ such that $\iota(s) = (i, \pi)$ for some $\pi \in \Pi$, of a proposal $\theta^s \in V$ and, for each state $s \in S$, of an acceptance set $A^{i,s} \subset V$. A stationary strategy of a player specifies a unique action for each of his decision nodes. This action depends only on the state and not on any other aspect of the history if the player is a proposer and on the state as well as the proposal made if the player is a responder. A stationary strategy profile (θ, A) leads to a unique probability distribution over payoffs in V, so determines the utility $u_i(\theta, A)$ of Player $i \in N$. Conditional utilities are denoted by $u_i(\theta, A \mid s)$. The *social acceptance set* in state $s \in S$ is defined as $A^s = \bigcap_{i \in N} A^{i,s}$. The social acceptance set consists of all alternatives that are unanimously accepted when proposed in state s.

Definition 2.1. *A stationary subgame perfect equilibrium (SSPE) is a profile of pure stationary strategies which is a subgame perfect equilibrium of the game.*

We make the following standard assumptions on V, where we use the notation $V_+ = V \cap \mathbb{R}^n_+$ and ∂V_+ is the set of weakly Pareto efficient points in V_+. Moreover, a vector η with $\|\eta\| = 1$ is said to be *normal* to the set V at a point $\bar{v} \in V$ if $(v - \bar{v})^\top \eta \leq 0$ for every $v \in V$. The set of all vectors η normal to V at $\bar{v}$ is called the *normal to* V at $\bar{v}$.

Assumption A The set V is closed, convex, and comprehensive from below. The origin lies in the interior of V. The set V_+ is bounded and all points in ∂V_+ are strongly Pareto efficient. There is a unique vector in the normal to V at every $v \in \partial V_+$.

A stationary strategy profile (θ, A) is said to have *no delay* if for every $s \in S$ it holds that $\theta^s \in A^s$. A stationary strategy profile (θ, A) is said to have *efficient proposals* if for every $s \in S$ it holds that $\theta^s \in \partial V_+$.

Apart from the analysis of $G(\delta)$, the literature also typically studies the behavior of equilibria when the continuation probability δ tends to 1.

Definition 2.2. The profile of proposals $\bar{\theta} = (\bar{\theta}^s)_{s \in S}$ is a *limit equilibrium* if there is a sequence $\{\delta_m\}_{m \in \mathbb{N}}$ of continuation probabilities in $[0, 1)$ converging to 1 and a sequence of profiles $\{\theta(\delta_m)\}_{m \in \mathbb{N}} = \{(\theta^s(\delta_m)_{s \in S})\}_{m \in \mathbb{N}}$, where $\theta(\delta_m)$ is an SSPE profile of proposals of the game $G(\delta_m)$, such that $\lim_{m \to \infty} \theta(\delta_m) = \bar{\theta}$.

A limit equilibrium is a profile of proposals that can be approximated arbitrarily close by an SSPE profile of proposals when the probability of breakdown is arbitrarily small. Of particular interest is the relationship between limit equilibria and the asymmetric Nash bargaining solution with positive weights $\mu \in \mathbb{R}^n_+ \setminus \{0\}$, denoted μ-ANBS and defined as follows.

Definition 2.3. The asymmetric Nash product *with weights $\mu \in \mathbb{R}^n_+ \setminus \{0\}$ is the function $f : V_+ \to \mathbb{R}$ defined by*

$$f(v) = \prod_{i \in N} (v_i)^{\mu_i}.$$

The μ-ANBS is the unique maximizer of the function f on the set V_+.

Britz, Herings, and Predtetchinski (2010) study the class of exogenous protocols characterized by the following assumption.

Assumption B It holds that $S = N$, for every $s \in S$, $\iota(s) = (s, \pi^0)$ with π^0 the identity, for all $j, k \in N$, $p^j = p^k$, and the matrix $M = [p^j(1), \ldots, p^j(n)]$ is irreducible.

An irreducible matrix M has a unique stationary distribution μ. Recall that a stationary distribution μ is a probability distribution on the set of states satisfying $M\mu = \mu$. The Markovian protocols satisfying Assumption B are sufficiently rich to encompass the alternating offer protocol, the rotating protocol and time-invariant probability protocol.

Britz, Herings, and Predtetchinski (2014) study the class of endogenous protocols characterized by the following assumption.

Assumption C It holds that $S = N \times \Pi$, for every $s \in S$, $\iota(s) = s$, for every $j \in N$, for all $s, s' \in S$, $p^j(s) = p^j(s')$. There exists $(i, \pi), (j, \pi') \in S$ such that $p^j(i, \pi)$ assigns positive probability to (j, π').

We associate to each protocol satisfying Assumption C the weights $\mu > 0$ given by $\mu_j = \sum_{(j,\pi)\in S} p^j_{(j,\pi)}(s)$, $j \in N$, where the choice of s is irrelevant by Assumption C, so μ_j is the probability that Player j becomes the next proposer conditional on a rejection. The class of protocols satisfying Assumption C is sufficiently rich to include the rejector-proposes protocol as well as the generalization by Kawamori (2008).

The following result follows from Britz, Herings, and Predtetchinski (2010, 2014).

Theorem 2.1. *If Assumptions A and B, or Assumptions A and C are satisfied, then*

1. *An SSPE exists.*

2. *Every SSPE has no delay.*

3. *Every SSPE has efficient proposals.*

4. *There is a unique limit equilibrium.*

5. *All limit equilibrium proposals are equal to the μ-ANBS.*

The characterization of limit equilibrium proposals as a weighted Nash bargaining solution has been shown in Binmore, Rubinstein, and Wolinsky (1986) for bilateral bargaining. For multilateral bargaining, this result is obtained in Hart and Mas-Colell (1996) for the time-invariant probability protocol with uniform recognition probabilities and in Miyakawa (2008) and Laruelle and Valenciano (2008) for general recognition probabilities. Kultti and Vartiainen (2010) derive this result for the rotating offer protocol. Theorem 2.1 includes these results as special cases. The five claims of Theorem 2.1 and in particular the limit equilibrium payoffs are independent of the order in which the responding players accept or reject the proposal. This is noteworthy in the case of an endogenous protocol. For instance, one might have conjectured that the rejector-proposes protocol favors the player who comes first in the responder order.

Assumption D below generalizes Assumptions B and C to the model considered in the present paper and, in particular, to the very general bargaining protocol introduced in our model description in Section 2.

Assumption D.

1. For every player $i \in N$, there are a state $s \in S$ and a responder order $\pi \in \Pi$ such that $\iota(s) = (i, \pi)$.

2. For every player $i \in N$, there is a pair of states s' and s'' in S such that $\iota(s'') = (i, \pi)$ for some $\pi \in \Pi$ and $p^i_{s''}(s') > 0$.

3. Given any pair (s', s'') of states in S, there exist a finite number $T \in \mathbb{N}$, a sequence of T players $i_0, i_1, \ldots, i_{T-1}$, and a sequence of $T+1$ states $s_0, s_1, \ldots, s_T$ such that $s_0 = s'$ and $s_T = s''$, and $p^{i_t}_{s_{t+1}}(s_t) > 0$ for all $t = 0, \ldots, T-1$.

The first part of the assumption says that every player is the proposer in some state. The second part says that for every player, there is one state such that if the player rejects a proposal in that state, he becomes the next proposer with strictly positive probability. The third part says that one can move from every state to every other state with positive probability in finitely many transitions.

In view of Theorem 2.1, one could conjecture that the five results stated in Theorem 2.1 hold in our model if Assumptions A and D are satisfied.[1] However, we will show in the sequel that this is not true. In particular, in Proposition 5.3 below, we will show that a bargaining protocol can satisfy Assumption D but still violate the no delay and efficient proposals properties of Theorem 2.1.

2.4 Failure of Properties 2, 3, 4, and 5

Ideally one would like to prove Theorem 2.1 for the entire class of bargaining protocols as laid down in Section 2, up to standard regularity assumptions as in Assumption A. In this section, we present an example where Properties 2, 3, 4, and 5 of Theorem 2.1 are violated. Moreover, the example is minimal in the following sense: It has $S = N = \{1, 2, 3\}$, and for two states out of three we have $p^j(s) = p^k(s)$, for all $j, k \in N$. With the exception of one state, the protocol therefore satisfies Assumption B and is exogenous.

Example 2.1. There are three players and three states, $S = N = \{1, 2, 3\}$. Each player is the proposer in one state and players respond in ascending order, so we have

$$\begin{aligned} \iota(1) &= (1, \pi^0), \\ \iota(2) &= (2, \pi^0), \\ \iota(3) &= (3, \pi^0), \end{aligned}$$

where π^0 is the identity. Players have to divide a surplus of one unit, $V = \{v \in \mathbb{R}^3 | v_1 + v_2 + v_3 \leq 1\}$. This set clearly satisfies Assumption A. In state $s = 1$, we follow the rejector-proposes protocol for Players 2 and 3. If Player 1 rejects his own proposal in state $s = 1$, then each of the three states is selected next with equal probability.

$$\begin{aligned} p^1(1) &= (\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}), \\ p^2(1) &= (0, 1, 0), \\ p^3(1) &= (0, 0, 1). \end{aligned}$$

States $s = 2, 3$ are absorbing.

$$\begin{aligned} p^i(2) &= (0, 1, 0), \quad i \in N, \\ p^i(3) &= (0, 0, 1), \quad i \in N. \end{aligned}$$

[1]We thank an anonymous referee for this suggestion.

□

The next result claims not only that equilibria may exhibit delay, but even makes the stronger statement that all SSPEs feature delay for sufficiently high values of δ.

Proposition 2.1. *For $\delta > 1/2$, every SSPE in Example 2.1 has delay.*

Proof. Suppose by way of contradiction that (θ, A) is an SSPE which has no delay. Consider a subgame starting with a proposal by Player 2 in state 2. In this subgame, Player 2 remains the proposer forever, and it is straightforward to verify that the subgame has a unique SSPE where Player 2 captures the entire surplus. It holds that $u(\theta, A \mid 2) = (0, 1, 0)$. By a completely symmetric argument, we find that $u(\theta, A \mid 3) = (0, 0, 1)$. Now consider a subgame starting with a proposal by Player 1 in state 1. Since (θ, A) has no delay, Player 2 accepts θ^1 and it holds that $\theta^1_2 \geq \delta$, since a rejection by Player 2 leads to a breakdown with probability $1 - \delta$ and a transition to state 2 and a payoff of 1 for Player 2 with probability δ. Similarly, it holds that $\theta^1_3 \geq \delta$, since a rejection by Player 3 leads to a breakdown with probability $1 - \delta$ and a transition to state 3 and a payoff of 1 for Player 3 with probability δ. It follows that

$$\theta^1_1 \leq 1 - \theta^1_2 - \theta^1_3 \leq 1 - 2\delta < 0.$$

Since Player 1 can ensure a non-negative payoff by a strategy that rejects all proposals, we have obtained a contradiction to (θ, A) being an SSPE which has no delay. □

The intuition behind the example is the following. If Player 2 rejects the proposal of Player 1, then the game goes to an absorbing state where Player 2 remains the proposer forever. It is well-known that in any SSPE of such a subgame, Player 2 would capture the entire surplus. Thus, when the game is in state 1 Player 2 can guarantee himself a payoff of δ by rejecting Player 1's proposal. In any SSPE with no delay, Player 1 would need to offer at least the amount δ to Player 2. The same argument applies to Player 3: If Player 3 rejects a proposal of Player 1, the game goes to an absorbing state where Player 3 remains the proposer forever and can capture the entire surplus. Thus, when Player 3 reacts to the proposal of Player 1, he will not accept any less than δ. Indeed, the sum of the responding players' reservation payoffs is equal to 2δ. We can see that if $\delta > \frac{1}{2}$, then the available surplus is not sufficient for Player 1 to pay the other two players their reservation payoffs. Consequently, if $\delta > \frac{1}{2}$, no agreement can be reached in state 1.

The next issue is whether there is an SSPE with delay in Example 2.1 for $\delta > 1/2$. Consider a strategy profile $(\bar{\theta}, \bar{A})$ with proposals $\bar{\theta}^1 \in V$, $\bar{\theta}^2 = (0, 1, 0)$, and $\bar{\theta}^3 = (0, 0, 1)$, and acceptance sets

$$\begin{array}{rcll} \bar{A}^{1,s} & = & \{v \in V \mid v_1 \geq 0\}, & s = 1,2,3, \\ \bar{A}^{2,1} & = & \{v \in V \mid v_2 \geq \delta,\ v_3 \geq \delta\}, & \\ \bar{A}^{2,2} & = & \{v \in V \mid v_2 \geq \delta\}, & \\ \bar{A}^{2,3} & = & \{v \in V \mid v_2 \geq 0\}, & \\ \bar{A}^{3,s} & = & \{v \in V \mid v_3 \geq \delta\}, & s = 1,3, \\ \bar{A}^{3,2} & = & \{v \in V \mid v_3 \geq 0\}. & \end{array}$$

When players play according to $(\bar{\theta}, \bar{A})$, Player 1 makes a particular proposal belonging to V in state 1, which will be rejected by some player when $\delta > 1/2$. More precisely, the proposal $\bar{\theta}^1$ in state 1 is rejected by Player 1 when $\bar{\theta}^1_1 < 0$ and is rejected by Player 2 otherwise. Note that in state 1, Player 2 would even reject the proposal $(0, 1, 0)$, since acceptance of such a proposal would lead to a rejection by Player 3, followed by breakdown of the negotiations or a transition to state 3.

Consider first the case where $\bar{\theta}^1_1 < 0$. Player 1 now rejects his own proposal, negotiations break down with probability $1 - \delta$ and continue with probability δ. If negotiations continue, transitions occur with equal probability to each of the three states, a rejection of proposal $\bar{\theta}^1$ by Player 1 in state 1, an acceptance of payoff vector $(0, 1, 0)$ in state 2, and an acceptance of payoff vector $(0, 0, 1)$ in state 3.

Consider next the situation where $\bar{\theta}^1_1 \geq 0$. Since $\bar{\theta}^1_1 \geq 0$ and $\delta > 1/2$, Player 2 rejects the proposal, negotiations break down with probability $1 - \delta$ and continue in state 2 with probability δ. In the latter case, the payoff vector $(0, 1, 0)$ is proposed and accepted.

Proposition 2.2. *For $\delta > 1/2$, the strategy profile $(\bar{\theta}, \bar{A})$ is an SSPE in Example 2.1.*

Proof. To show that $(\bar{\theta}, \bar{A})$ is an SSPE, it suffices to verify the one-shot deviation property, see for instance Fudenberg and Tirole (1991). We consider three cases, depending on the state to which a decision node belongs.

Case 1. Decision nodes in state 1.
After a history in state 1 where Player 1 has to propose, the proposal $\bar{\theta}^1$ is rejected, either by Player 1 in case $\bar{\theta}^1_1 < 0$ or by Player 2 in case $\bar{\theta}^1_1 \geq 0$, and leads ultimately to breakdown, or the acceptance of proposal $\bar{\theta}^2$, or the acceptance of proposal $\bar{\theta}^3$. In all cases, Player 1 receives a payoff of zero. A one-shot deviation to any other proposal is rejected as well, either by Player 1 or by Player 2, and also leads ultimately to a payoff of zero for sure. Such a deviation is therefore not profitable.

After a history in state 1 where Player 1 has to respond, any proposal v with $v_1 < 0$ is rejected by Player 1, and ultimately leads to a payoff of zero for sure. A one-shot deviation to acceptance leads to the acceptance of v and a negative payoff for Player 1, or the rejection of v by Player 2 or Player 3 and a payoff of zero for Player 1. Such a deviation is therefore not profitable. Any proposal v with $v_1 \geq 0$ is accepted by Player 1, next rejected by Player 2, and followed by breakdown of the negotiations or acceptance of $(0, 1, 0)$ in the next period. The payoff for Player 1 is therefore zero. A one-shot deviation to rejection leads ultimately to a payoff of zero for Player 1 as well and is therefore not profitable.

After a history in state 1 where Player 2 has to respond, any proposal v with $v_2 < \delta$ or $v_3 < \delta$ is rejected by Player 2, which results in a payoff of δ for Player 2. A one-shot deviation to acceptance is followed by an acceptance by Player 3 if $v_3 \geq \delta$ and leads to payoff $v_2 < \delta$ for Player 2, so is not profitable, and is followed by a rejection by Player 3 if $v_3 < \delta$, leading to payoff 0 for Player 2, so is not profitable either. Any proposal v with $v_2 \geq \delta$ and $v_3 \geq \delta$ is accepted by Player 2, followed by an acceptance by Player 3, and a payoff of v_2 for Player 2. A one-shot deviation to rejection leads to a payoff of $\delta \leq v_2$ for Player 2 and is therefore not profitable.

After a history in state 1 where Player 3 has to respond, any proposal v with $v_3 < \delta$ is rejected by Player 3, resulting in a payoff of δ for Player 3. A one-shot deviation to acceptance is clearly not profitable. Any proposal v with $v_3 \geq \delta$ is accepted by Player 3, leading to a payoff of v_3 for Player 3. A one-shot deviation to rejection is clearly not profitable.

Case 2. Decision nodes in state 2.
After a history in state 2 where Player 2 has to propose, the proposal $\bar{\theta}^2 = (0, 1, 0)$ by Player 2 is accepted by all players, and leads to utility 1 for Player 2. Since Players 1 and 3 reject proposals which give them a negative payoff, there are no profitable one-shot deviations for Player 2. Since a one-shot rejection by any player leads to payoffs $(0, \delta, 0)$, the one-shot deviation property holds for responders.

Case 3. Decision nodes in state 3.
This case is similar to Case 2. □

Proposition 4.3 describes a continuum of SSPEs, parametrized by the proposal $\bar{\theta}^1$ by Player 1. For equilibria with $\bar{\theta}^1_1 < 0$, the equilibrium payoffs when starting in state 1 are equal to $u(\bar{\theta}, \bar{A} \mid 1) = (0, \delta/(3 - \delta), \delta/(3 - \delta))$. For equilibria with $\bar{\theta}^1_1 \geq 0$, it holds that $u(\bar{\theta}, \bar{A} \mid 1) = (0, \delta, 0)$. None of the properties, apart from SSPE existence, mentioned in Theorem 2.1 are satisfied. All SSPEs have delay. There is a continuum of SSPEs where Player 1 makes an inefficient proposal, and even if Player 1 makes an efficient proposal, it is still rejected by Player 1 or Player 2. Any element of $V \times \{(0, 1, 0)\} \times \{(0, 0, 1)\}$ can be a limit equilibrium proposal, so there is no unique limit equilibrium. When starting in state 1, limit equilibrium utilities are either equal to $(0, 1/2, 1/2)$ or $(0, 0, 1)$. Finally, limit equilibrium proposals are not equal to each other.

2.5 Robustness of the Example

In this section, we will examine the robustness of Example 2.1 to perturbations of the transition probabilities. In particular, we will see that the presence of absorbing states is not vital for equilibrium delay, inefficient proposals, multiple limit equilibria, and limit equilibrium proposals that are not equal to each other.

Indeed, one may object to Example 2.1 that states 2 and 3 are absorbing, no matter what actions the players take, and therefore Assumption D is violated. Still, we would like to argue that Example 2.1 is robust to perturbations in the transition probabilities and that the violations of Properties 2, 3, 4, and 5 of Theorem 2.1 are not due to the presence of absorbing states. Indeed, we will now present a perturbed version of Example 2.1 which satisfies Assumption D and show that nevertheless the Properties 2, 3, 4, and 5 of Theorem 2.1 are violated. To be more specific, consider the case where all the transition probabilities are perturbed by some $\varepsilon \in (0, 1/3)$. We obtain the following example.

Example 2.2. There are three players and three states, $S = N = \{1, 2, 3\}$. Each player is the proposer in one state and players respond in ascending order, so we have

$$\begin{array}{rcl} \iota(1) & = & (1, \pi^0), \\ \iota(2) & = & (2, \pi^0), \\ \iota(3) & = & (3, \pi^0), \end{array}$$

where π^0 is the identity. Players have to divide a surplus of one unit, $V = \{v \in \mathbb{R}^3 | v_1 + v_2 + v_3 \leq 1\}$. In state $s = 1$, the transitions depend on the identity of the player who rejects a proposal. If Player 1 rejects his own proposal in state $s = 1$, then each of the three states is selected next with equal probability. If Player $j = 2, 3$ rejects the proposal by Player 1, then we move to the state in which j is proposer with probability $1 - 2\varepsilon$. If we do not go to the state in which player j proposes, then one of the other two states is chosen randomly.

$$\begin{array}{rcl} p^1(1) & = & (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}), \\ p^2(1) & = & (\varepsilon, 1 - 2\varepsilon, \varepsilon), \\ p^3(1) & = & (\varepsilon, \varepsilon, 1 - 2\varepsilon). \end{array}$$

In states $s = 2, 3$, the transitions are independent of the identity of the rejecting player. The process returns to the same state with probability $1 - 2\varepsilon$. If the process does not return to the same state, then one of the other two states is chosen at random.

$$\begin{array}{rcl} p^i(2) & = & (\varepsilon, 1 - 2\varepsilon, \varepsilon), \quad i \in N, \\ p^i(3) & = & (\varepsilon, \varepsilon, 1 - 2\varepsilon), \quad i \in N. \end{array}$$

□

Note that the transition dynamics in this example satisfy Assumption D. In particular, Assumption D.1 is satisfied because each player is the proposer in one of the three states. Moreover, the example is such that regardless of the current state and the identity of the rejector, we move with with positive probability to every state after every rejection. Thus, Assumptions D.2 and D.3 are also satisfied.

Consider a strategy profile $(\hat{\theta}, \hat{A})$ with proposals $\hat{\theta}^1 \in \{v \in V \mid v_1 < 0\}$,

$$\begin{aligned} \hat{\theta}^2 &= (0, 1-y, y), \\ \hat{\theta}^3 &= (0, y, 1-y), \end{aligned}$$

where

$$y = \frac{3\delta\varepsilon}{(3-\delta)(1-\delta+3\delta\varepsilon)},$$

and acceptance sets

$$\begin{aligned} \hat{A}^{1,s} &= \{v \in V \mid v_1 \geq 0\}, & s = 1,2,3, \\ \hat{A}^{2,1} &= \{v \in V \mid v_2 \geq z,\ v_3 \geq z\}, \\ \hat{A}^{2,2} &= \{v \in V \mid v_2 \geq z\}, \\ \hat{A}^{2,3} &= \{v \in V \mid v_2 \geq y\}, \\ \hat{A}^{3,s} &= \{v \in V \mid v_3 \geq z\}, & s = 1,3, \\ \hat{A}^{3,2} &= \{v \in V \mid v_3 \geq y\}, \end{aligned}$$

where

$$z = \frac{\delta(1-\delta)(3-\delta-9\delta\varepsilon^2) + \delta\varepsilon(2-\delta)(6\delta-3)}{(3-\delta)(1-\delta+3\delta\varepsilon)}.$$

Note that if ε tends to zero, then y tends to zero and z to δ.

When players play according to $(\hat{\theta}, \hat{A})$, Player 1 makes a proposal $\hat{\theta}^1$ in V with $\hat{\theta}^1_1 < 0$ in state 1, which is rejected by Player 1 himself, and a transition to each of the three states follows with equal probability. In state 2, Player 2 makes a proposal that gives a payoff of 0 to Player 1, gives the reservation payoff y to Player 3, and keeps the remainder of the surplus himself. State 3 is symmetric to state 2, with the roles of Players 2 and 3 reversed. The proposals $\hat{\theta}^2$ and $\hat{\theta}^3$ in states 2 and 3 are accepted since $1 - y > z$, which follows from the fact that

$$(3-\delta)(1-\delta) + 3\delta\varepsilon(2-\delta) > \delta(1-\delta)(3-\delta-9\delta\varepsilon^2) + \delta\varepsilon(2-\delta)(6\delta-3).$$

Proposition 2.3. *For every $\varepsilon \in (0, 1/6)$, there exists $\bar{\delta} < 1$ such that for every $\delta \geq \bar{\delta}$ the strategy profile $(\hat{\theta}, \hat{A})$ is an SSPE in Example 2.2.*

Proof. For $s = 1, 2, 3$, we define the equilibrium utilities conditional on state s, $x^s = u(\hat{\theta}, \hat{A} \mid s)$. The symmetry of the game and the strategies implies that $x^1_2 = x^1_3$,

$x_1^2 = x_1^3$, $x_2^2 = x_3^3$, and $x_3^2 = x_2^3$. It holds that $x^1 = (0, \delta/(3-\delta), \delta/(3-\delta))$, $x^2 = \hat{\theta}^2$, and $x^3 = \hat{\theta}^3$, where the expression for x^1 uses the observation that $x_2^1 = (\delta/3) + (\delta/3)x_2^1$.

To show that $(\hat{\theta}, \hat{A})$ is an SSPE, we verify the one-shot deviation property. We consider three cases, depending on the state to which a decision node belongs.

Case 1. Decision nodes in state 1.
Consider a history in state 1 after which Player 3 has to respond. A rejection followed by play according to $(\hat{\theta}, \hat{A})$ leads to a payoff for Player 3 equal to $\delta\varepsilon x_3^1 + \delta\varepsilon x_3^2 + \delta(1-2\varepsilon)x_3^3$. A straightforward, but tedious, calculation reveals that this payoff is equal to z. Since Player 3 accepts proposals in state 1 if and only if $v_3 \geq z$, this shows that the one-shot deviation property is satisfied.

Consider a history in state 1 after which Player 2 has to respond to a proposal v. Suppose first that $v_3 \geq z$. A calculation similar to that in the previous paragraph shows that rejection of v by Player 2 yields a payoff of z. Acceptance of v by Player 2 leads to the payoff v_2 since v is also accepted by Player 3. We conclude that accepting the proposal v if and only if $v_2 \geq z$ does not violate the one-shot deviation principle.

Suppose now that $v_3 < z$. As before, rejecting v yields Player 2 a payoff of z. If Player 2 accepts v, it is rejected by Player 3, yielding a continuation utility equal to $\delta\varepsilon x_2^1 + \delta\varepsilon x_2^2 + \delta(1-2\varepsilon)x_2^3$. A straightforward, though tedious, calculation reveals the latter expression to be equal to y, and it holds that $y < z$ since

$$z - y = \frac{\delta(1-\delta)(1-3\varepsilon)(3-\delta+3\delta\varepsilon)}{(3-\delta)(1-\delta+3\delta\varepsilon)} > 0.$$

Hence v should be rejected by Player 2.

The verification of the one-shot deviation property for Player 1 is trivial for histories where he responds. Consider a history where Player 1 proposes. Player 1 has a profitable one-shot deviation if and only if he can make a proposal that gives more than z to Players 2 and 3 and a positive payoff to himself. This implies that Player 1 has no profitable one-shot deviation if $2z \geq 1$. This inequality is satisfied if and only if $\delta \geq \frac{1}{2(1-3\varepsilon)}$. It follows that for every $\varepsilon \in (0, 1/6)$, there exists $\bar{\delta} < 1$ such that for every $\delta \geq \bar{\delta}$, $2z \geq 1$.

Case 2. Decision nodes in state 2.
Player 3 accepts a proposal v if and only if $v_3 \geq y$, where y equals the continuation payoff of Player 3 following his rejection. This shows that the one-shot deviation property is satisfied.

Player 2 accepts a proposal v if and only if $v_2 \geq z$, where z equals the continuation payoff of Player 2 following his rejection. We observe that acceptance of v yields Player 2 payoff v_2 if Player 3 accepts as well, and z if Player 3 rejects v. This shows that the one-shot deviation property is satisfied.

The verification of the one-shot deviation property for Player 1 is trivial.

Consider a history in state 2 after which Player 2 proposes. Since the proposal $\hat{\theta}^2$ of Player 2 gives Players 1 and 3 the least amount they are willing to accept, there is no profitable one-shot deviation for Player 2 which will be accepted by Players 1 and 3. Consider a one-shot deviation by Player 2 which is rejected by some player. Ultimately, such a proposal leads to breakdown and payoff 0 for Player 2, or an acceptance of $\hat{\theta}^2$ and payoff $1-y$ for Player 2 or an acceptance of $\hat{\theta}^3$ and Payoff y for Player 2. Since $y < 1-y$, the expected payoff for Player 2 is less than $x_2^2 = 1-y$, so the deviation is not profitable.

Case 3. Decision nodes in state 3.
By symmetry, the line of argument is the same as in Case 2. □

The equilibrium presented in Proposition 2.3 clearly violates Properties 2 and 3 of Theorem 2.1: No agreement is reached in state 1, and the proposal made in state 1 need not be efficient. The equilibrium also has the remarkable feature that Player 1's equilibrium payoff is equal to zero in every state. This is surprising, since in the framework of Britz, Herings, and Predtetchinski (2014), the bargaining power of a player is proportional to the probability to propose conditional on his own rejection. In the protocol of Example 2.2, this probability is at least ε for Player 1. Proposition 5.2 also gives us a multitude of limit equilibria, so that properties 4 and 5 in Theorem 2.1 are also violated. More specifically, we can see that any point $(\theta^1, \theta^2, \theta^3)$ where $\theta^1 \in V$ with $\theta^1_1 < 0$ and $\theta^2 = \theta^3 = (0, \frac{1}{2}, \frac{1}{2})$ is a limit equilibrium.

Consider a strategy profile $(\tilde{\theta}, \tilde{A})$ with proposals $\tilde{\theta}^1 \in \{v \in V \mid v_1 \geq 0\}$,

$$\begin{aligned}\tilde{\theta}^2 &= (0, 1-y_3, y_3),\\ \tilde{\theta}^3 &= (0, y_2, 1-y_2),\end{aligned}$$

where

$$\begin{aligned}y_2 &= \frac{\delta\varepsilon + \delta^2\varepsilon - 3\delta^2\varepsilon^2}{1 - \delta + 2\delta\varepsilon + \delta^2\varepsilon - 3\delta^2\varepsilon^2},\\ y_3 &= \frac{\delta\varepsilon}{1 - \delta + 2\delta\varepsilon + \delta^2\varepsilon - 3\delta^2\varepsilon^2},\end{aligned}$$

and acceptance sets

$$
\begin{aligned}
\tilde{A}^{1,s} &= \{v \in V \mid v_1 \geq 0\}, && s = 1,2,3,\\
\tilde{A}^{2,1} &= \{v \in V \mid v_2 \geq z_2,\ v_3 \geq z_3\}, &&\\
\tilde{A}^{2,2} &= \{v \in V \mid v_2 \geq z_2\}, &&\\
\tilde{A}^{2,3} &= \{v \in V \mid v_2 \geq y_2\}, &&\\
\tilde{A}^{3,s} &= \{v \in V \mid v_3 \geq z_3\}, && s = 1,3,\\
\tilde{A}^{3,2} &= \{v \in V \mid v_3 \geq y_3\}, &&
\end{aligned}
$$

where

$$
\begin{aligned}
z_2 &= \frac{\delta - \delta^2 - 2\delta\varepsilon + 4\delta^2\varepsilon - 3\delta^2\varepsilon^2}{1 - \delta + 2\delta\varepsilon + \delta^2\varepsilon - 3\delta^2\varepsilon^2},\\
z_3 &= \frac{\delta - \delta^2 - 2\delta\varepsilon + 3\delta^2\varepsilon}{1 - \delta + 2\delta\varepsilon + \delta^2\varepsilon - 3\delta^2\varepsilon^2}.
\end{aligned}
$$

Note that if ε tends to zero, then y_2 and y_3 tend to zero and z_2 and z_3 tend to δ.

When players play according to $(\tilde{\theta}, \tilde{A})$, Player 1 makes a proposal $\tilde{\theta}^1$ in V with $\tilde{\theta}^1_1 \geq 0$ in state 1. A straightforward calculation shows that $z_2 + z_3 > 1$ if and only if

$$(1 - \delta)(2\delta - 6\delta\varepsilon - 1) > 0.$$

Therefore it holds that if $\varepsilon < 1/6$ and $\delta > 1/(2 - 6\varepsilon)$, then Player 2 rejects $\tilde{\theta}^1$, and a transition to state 2 follows with high probability. In state 2, Player 2 makes a proposal that gives a payoff of 0 to Player 1, gives the reservation payoff y_2 to Player 3, and keeps the remainder of the surplus himself. State 3 is similar, with the roles of Players 2 and 3 reversed. The proposals $\tilde{\theta}^2$ and $\tilde{\theta}^3$ in states 2 and 3 are accepted since $1 - y_3 \geq z_2$ and $1 - y_2 \geq z_3$ which follows respectively from

$$
\begin{aligned}
&1 - \delta + \delta\varepsilon + \delta^2\varepsilon - 3\delta^2\varepsilon^2 > \delta(1 - \delta) + 2\delta^2\varepsilon - 3\delta^2\varepsilon^2 > \delta - \delta^2 - 2\delta\varepsilon + 4\delta^2\varepsilon - 3\delta^2\varepsilon^2,\\
&1 - \delta + \delta\varepsilon > \delta - \delta^2 - 2\delta\varepsilon + 3\delta^2\varepsilon.
\end{aligned}
$$

Proposition 2.4. *For every $\varepsilon \in (0, 1/6)$, there exists $\bar{\delta} < 1$ such that for every $\delta \geq \bar{\delta}$ the strategy profile $(\tilde{\theta}, \tilde{A})$ is an SSPE in Example 2.2.*

Proof. For $s = 1, 2, 3$, we define the equilibrium utilities conditional on state s, $x^s = u(\tilde{\theta}, \tilde{A} \mid s)$. We have that $x^2 = \tilde{\theta}^2$, $x^3 = \tilde{\theta}^3$, $x^1_1 = 0$, $x^1_2 = \delta\varepsilon x^1_2 + \delta(1 - 2\varepsilon)x^2_2 + \delta\varepsilon x^3_2$, so

$$x^1_2 = \frac{\delta - \delta^2 - 2\delta\varepsilon + 4\delta^2\varepsilon - 3\delta^2\varepsilon^2}{1 - \delta + 2\delta\varepsilon + \delta^2\varepsilon - 3\delta^2\varepsilon^2},$$

and $x^1_2 = x^2_3$, since a rejection by Player 2 in state 1 leads to the same transitions as a rejection by Player 3 in state 2.

To show that $(\tilde{\theta}, \tilde{A})$ is an SSPE, we verify the one-shot deviation property. We consider three cases, depending on the state to which a decision node belongs.

Case 1. Decision nodes in state 1.
The following argument is analogous to the argument in the proof of the previous proposition, except that z_2, z_3 now differ.

Consider a history in state 1 after which Player 3 has to respond. A rejection followed by play according to $(\tilde{\theta}, \tilde{A})$ leads to a payoff for Player 3 equal to $\delta\varepsilon x_3^1 + \delta\varepsilon x_3^2 + \delta(1-2\varepsilon)x_3^3 = z_3$. Since Player 3 accepts proposals in state 1 if and only if $v_3 \geq z_3$, this shows that the one-shot deviation property is satisfied.

Consider a history in state 1 after which Player 2 has to respond to a proposal v. Suppose first that $v_3 \geq z_3$. A calculation similar to that in the previous paragraph shows that rejection of v by Player 2 yields Player 2 a payoff of z_2. Acceptance yields v_2 because v is accepted by Player 3. Hence accepting v if and only if $v_2 \geq z_2$ does not violate the one-shot deviation principle.

Suppose now that $v_3 < z_3$. As before, rejecting v by Player 2 gives payoff z_2. If Player 2 accepts v, then it is rejected by Player 3 and yields $\delta\varepsilon x_2^1 + \delta\varepsilon x_2^2 + \delta(1-2\varepsilon)x_2^3 = y_2$. It holds that $y_2 < z_2$ since

$$z_2 - y_2 = \frac{\delta(1-\delta)(1-3\varepsilon)}{1 - \delta + 2\delta\varepsilon + \delta^2\varepsilon - 3\delta^2\varepsilon^2} > 0.$$

Thus rejecting v does not violate the one-shot deviation principle.

The verification of the one-shot deviation property for Player 1 is trivial for histories where he responds. We have already argued that $z_2 + z_3 > 1$ if $\delta > 1/(2-6\varepsilon)$. For such values of δ, Player 1 cannot make a profitable one-shot deviation as a proposer.

Case 2. Decision nodes in state 2.
Using the same argument as in the proof of the previous proposition, we can show that the one-shot deviation property is satisfied.

Consider a history in state 2 after which Player 2 proposes. Since the proposal $\tilde{\theta}^2$ of Player 2 gives Players 1 and 3 the least amount they are willing to accept, there is no profitable one-shot deviation for Player 2 which will be accepted by Players 1 and 3. Consider a one-shot deviation by Player 2 which is rejected by some player. Ultimately, such a proposal leads to breakdown and payoff 0 for Player 2, or an acceptance of $\tilde{\theta}^2$ and payoff $1 - y_3$ for Player 2 or an acceptance of $\tilde{\theta}^3$ and Payoff y_2 for Player 2. Since it is easily verified that $y_2 < 1 - y_3$, the expected payoff for Player 2 is less than $x_2^2 = 1 - y_3$, so the deviation is not profitable.

Case 3. Decision nodes in state 3.
This case is analogous to the Case 2.

□

Note that the strategy profile $(\tilde{\theta}, \tilde{A})$ violates no delay and efficient proposals which are two of the properties listed in Theorem 2.1. This is because there is no agreement in state 1, and because the proposal $\tilde{\theta}^1$, which is indeterminate, may be inefficient. Proposition 5.3 above yields limit equilibria of the form $(\theta^1, \theta^2, \theta^3)$ where $\theta^1 \in V$ with $\theta^1_1 \geq 0$ and

$$\theta^2 = \theta^3 = \left(0, \frac{2-3\varepsilon}{3-3\varepsilon}, \frac{1}{3-3\varepsilon}\right).$$

Clearly, not all players make the same proposals in the limit. The fact that in both cases Players 2 and 3 do make the same proposals in the limit can be deduced from Theorem 3.4. Indeed, since Player 1's proposal is rejected, one can view the resulting system, involving Players 2 and 3 only, as an exogenous protocol.

2.6 Non-existence of SSPEs

In the example of the previous section, all properties of Theorem 2.1 are violated, with the exception of the existence of an SSPE. In this section, we will present an example where no SSPE exists at all, neither one with immediate agreement, nor one with delay. The only modification when compared to Example 2.1 is that in state 1, after a rejection by Player 2 or 3, we follow the rejector-proposes protocol with probability 1/2 and return to state 1 with the complementary probability.

Example 2.3. There are three players and three states, $S = N = \{1, 2, 3\}$. Each player is the proposer in one state and players respond in ascending order, so we have

$$\begin{aligned} \iota(1) &= (1, \pi^0), \\ \iota(2) &= (2, \pi^0), \\ \iota(3) &= (3, \pi^0), \end{aligned}$$

where π^0 is the identity. Players have to divide a surplus of one unit, $V = \{v \in \mathbb{R}^3 \mid v_1 + v_2 + v_3 \leq 1\}$. In state $s = 1$, the transitions depend on the identity of the player who rejects a proposal,

$$\begin{aligned} p^1(1) &= \left(\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}\right), \\ p^2(1) &= \left(\tfrac{1}{2}, \tfrac{1}{2}, 0\right), \\ p^3(1) &= \left(\tfrac{1}{2}, 0, \tfrac{1}{2}\right). \end{aligned}$$

States $s = 2, 3$ are absorbing.

$$\begin{aligned} p^i(2) &= (0, 1, 0), \quad i \in N, \\ p^i(3) &= (0, 0, 1), \quad i \in N. \end{aligned}$$

$\square$

Proposition 2.5. *For $\delta \in (2/3, 3/4)$, there is no SSPE in Example 2.3.*

Proof. Suppose (θ, A) is an SSPE. It clearly holds that $\theta^2 = (0,1,0)$ and $\theta^3 = (0,0,1)$. Now let $x = u(\theta, A \mid 1)$ be the equilibrium utilities conditional on state 1 and let z^i be the vector of continuation payoffs after a proposal in state 1 is rejected by Player i. We have

$$\begin{aligned} z^1 &= \tfrac{\delta}{3}x + \tfrac{\delta}{3}(0,1,0) + \tfrac{\delta}{3}(0,0,1), \\ z^2 &= \tfrac{\delta}{2}x + \tfrac{\delta}{2}(0,1,0), \\ z^3 &= \tfrac{\delta}{2}x + \tfrac{\delta}{2}(0,0,1). \end{aligned}$$

We distinguish four possible cases.

Case 1. θ^1 is accepted by Players 1, 2, and 3.
In this case $x = \theta^1$. Since all players accept θ^1, we have

$$\begin{aligned} x_1 &\geq z^1_1 = \tfrac{\delta}{3}x_1, \\ x_2 &\geq z^2_2 = \tfrac{\delta}{2}x_2 + \tfrac{\delta}{2}, \\ x_3 &\geq z^3_3 = \tfrac{\delta}{2}x_3 + \tfrac{\delta}{2}. \end{aligned}$$

This gives the inequalities $x_1 \geq 0$, $x_2 \geq \delta/(2-\delta)$, and $x_3 \geq \delta/(2-\delta)$, whereas at the same time $x_1 + x_2 + x_3 = \theta^1_1 + \theta^1_2 + \theta^1_3 \leq 1$. For $\delta > 2/3$, we have a contradiction.

Case 2. θ^1 is accepted by Players 1 and 2, and rejected by Player 3.
Since Player 3 rejects θ^1, we have $x = z^3$, and in particular, $x_2 = z^3_2 = (\delta/2)x_2$. Since $\delta/2 < 1$, it follows that $z^3_2 = x_2 = 0$. But Player 2 accepts, so it must hold that $z^3_2 = 0 \geq z^2_2 = \delta/2$, contradicting that δ is strictly positive.

Case 3. θ^1 is accepted by Player 1 and rejected by Player 2.
Since Player 2 rejects θ^1, we have $x = z^2$, from which it follows that $x = (0, \delta/(2-\delta), 0)$ and in particular that $z^2_2 = \delta/(2-\delta)$. Plugging $x_1 = 0$ into the equation for z^1 and $x_3 = 0$ into the equation for z^3 gives $z^1_1 = 0$ and $z^3_3 = \delta/2$, respectively. Since $z^1_1 + z^2_2 + z^3_3 = 0 + \delta/2 + \delta/(2-\delta) < 1$ if $\delta < 3/4$, there exists a proposal $v \in V$ such that $v \gg (z^1_1, z^2_2, z^3_3)$. Clearly, v would be accepted by all players. Since $v_1 > x_1 = 0$, proposing v instead of θ^1 would be a profitable deviation for Player 1.

Case 4. θ^1 is rejected by Player 1.
If Player 1 rejects his own proposal, then $x = z^1$, so $x = (0, \delta/(3-\delta), \delta/(3-\delta))$, and in particular $z^1_1 = 0$. Plugging the expression for x into the equations for z^2 and z^3, we find that $z^2_2 = z^3_3 = 3\delta/(6-2\delta)$. Given that $\delta < 3/4$, we have that $z^1_1 + z^2_2 + z^3_3 < 1$, thus there exists a proposal v such that $v \in V$ and $v \gg (z^1_1, z^2_2, z^3_3)$. The proposal v would be accepted unanimously. Since $v_1 > x_1 = 0$, it would be a profitable deviation for Player 1

to propose v instead of θ^1. □

Perturbing the transition probabilities slightly will not affect the conclusion of the proposition. One can show that the set of transition probabilities such that the resulting bargaining game has an SSPE is closed.

2.7 Conclusion

We consider bargaining games of perfect information with a unanimous acceptance rule. The focus of the analysis is on the influence of the bargaining protocol on the bargaining outcomes. We consider a framewrok where the proposer and the order of the responding players is determined by a state variable. The probability distribution over states in the following period is determined jointly by the current state and the identity of the player who rejected the previous proposal.

Many papers in the existing literature are incorporated as special cases of this framework. It is either assumed that the identity of the future proposer only depends on the identity of the current proposer or it is assumed that it only depends on the identity of the rejector. In both cases it holds that SSPEs exist, have the immediate acceptance property, and involve efficient proposals only. Asymptotically, as players become perfectly patient, all such equilibria converge to an appropriately defined asymmetric Nash bargaining solution.

In this paper, however, we show that these conclusions do not carry over to the general framework: SSPEs need not exist. When they do exist, they may exhibit delay and involve inefficient proposals. The limit equilibrium need not be unique, limit equilibrium utilities need not be unique, and players may make different equilibrium proposals in the limit.

The central message in Britz, Herings, and Predtetchinski (2014) is that the bargaining power of a player is determined by the probability to propose conditional on his own rejection. For the more general framework considered here, we show that even though players can propose with positive probability conditional on their own rejection, they might still have no bargaining power at all.

References

Ausubel, L.M., P. Cramton, and R.J. Deneckere (2002), "Bargaining with Incomplete Information," in R.J. Aumann and S. Hart (eds.), *Handbook of Game Theory, Volume 3,* Elsevier Science Publishers, Amsterdam, pp. 1897–1945.

Banks, J., and J. Duggan (2000), "A Bargaining Model of Collective Choice," *American Political Science Review, 94,* 73–88.

Binmore, K. (1987), "Perfect Equilibria in Bargaining Models," in K. Binmore and P. Dasgupta (eds.), *The Economics of Bargaining,* Basil Blackwell, Oxford, UK, pp. 77–105.

Binmore, K., A. Rubinstein, and A. Wolinsky (1986), "The Nash Bargaining Solution in Economic Modelling," *RAND Journal of Economics, 17,* 176–188.

Bloch, F. (1996), "Sequential Formation of Coalitions in Games with Externalities and Fixed Payoff Division," *Games and Economic Behavior, 14,* 90–123.

Britz, V., P.J.J. Herings, and A. Predtetchinski (2010), "Non-cooperative Support for the Asymmetric Nash Bargaining Solution," *Journal of Economic Theory, 145,* 1951–1967.

Britz, V., P.J.J. Herings, and A. Predtetchinski (2014), "On the Convergence to the Nash Bargaining Solution for Action-dependent Bargaining Protocols," *Games and Economic Behavior, 86,* 178–183.

Chatterjee, K., B. Dutta, D. Ray, and K. Sengupta (1993), "A Noncooperative Theory of Coalitional Bargaining," *Review of Economic Studies, 60,* 463–477.

Fudenberg, D., and J. Tirole (1991), *Game Theory,* MIT Press, Cambridge, Massachusetts.

Haller, H. (1986), "Non-cooperative Bargaining of $N \geq 3$ Players," *Economics Letters, 22,* 11–13.

Hart, S., and A. Mas-Colell (1996), "Bargaining and Value," *Econometrica, 64,* 357–380.

Herings, P.J.J., and A. Predtetchinski (2009), "Bargaining with Non-convexities," METEOR Research Memorandum 09/42, Maastricht University, Maastricht, pp. 1–16.

Herrero, M.J. (1985), *A Strategic Bargaining Approach to Market Institutions,* PhD Thesis, London School of Economics, 1–111.

Jéhiel, P. and B. Moldovanu (1995), "Cyclical Delay in Bargaining with Externalities," *The Review of Economic Studies*, 62, 619–637.

Kalandrakis, T. (2004), "Equilibria in Sequential Bargaining Games as Solutions to Systems of Equations," *Economics Letters, 84,* 407–411.

Kawamori, T. (2008), "A Note on Selection of Proposers in Coalitional Bargaining," *International Journal of Game Theory, 37,* 525–532.

Kultti, K., and H. Vartiainen (2010), "Multilateral Non-Cooperative Bargaining in a General Utility Space," *International Journal of Game Theory, 39,* 677–689.

Laruelle, A., and F. Valenciano (2008), "Noncooperative Foundations of Bargaining Power in Committees and the Shapley-Shubik Index," *Games and Economic Behavior, 63,* 341–353.

Maskin, E.S., and J. Tirole (2001), "Markov Perfect Equilibrium, I. Observable Actions," *Journal of Economic Theory, 100,* 191–219.

Merlo, A. and C.Wilson (1995), "A Stochastic Model of Sequential Bargaining with Complete Information," *Econometrica, 63,* 371–399.

Miyakawa, T. (2008), "Noncooperative Foundation of n-Person Asymmetric Nash Bargaining Solution," *Journal of Economics of Kwansei Gakuin University, 62,* 1–18.

Rubinstein, A. (1982), "Perfect Equilibrium in a Bargaining Model," *Econometrica, 50,* 97–109.

Selten, R. (1981), "A Noncooperative Model of Characteristic Function Bargaining," in V. Böhm and H.H. Nachtkamp (eds.), *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern,* Bibliografisches Institut Mannheim, pp. 131–151.

Yildiz, M. (2003), "Bargaining without a Common Prior: An Immediate Agreement Theorem," *Econometrica, 71,* 793–811.

Chapter 3

Rent-Seeking and Surplus Destruction in Unanimity Bargaining Games

Published in *Games and Economic Behavior*

Abstract: In non–cooperative bargaining games in the tradition of Rubinstein, the proposer's bargaining power stems from the prospect of a delay in case of disagreement. Since players are impatient, this delay is costly for everyone. We consider a unanimity bargaining game in which the proposer can strategically choose the length of this delay. We assume that the size of the surplus depends endogenously on the chosen length of the prospective delay. Intuitively, the proposer faces the following trade–off: The more he exploits his proposer power, the smaller is the surplus that can be divided. One interpretation is that aggressive bargaining tactics hurt the fruitful cooperation among players, and thus the surplus. We characterize stationary equilibrium strategies and payoffs, and obtain sharp predictions on the extent of surplus destruction, the size of the social loss, and the surplus allocation.

JEL classification: C72, C78
Keywords: Bargaining, Rent-seeking, Surplus destruction, Discount factor, Timing, Commitment

Acknowledgements: The author would like to thank Hans Gersbach, Faruk Gul, Matthew Jackson, and Santiago Oliveiros, as well as the audience at the 2017 European Meeting on Game Theory in Paris for helpful comments.

3.1 Introduction

Unanimity bargaining games in the tradition of Rubinstein (1982) are widely used in economics and political science to study negotiations on the division of a surplus. The bargaining process is modeled as a sequence of rounds. In each round, one player makes a proposal. Unless this proposal is unanimously accepted, bargaining proceeds to another round. Every unsuccessful round comes at a social cost of disagreement ("bargaining friction") which is typically modeled as a delay. The proposer's bargaining power increases with this social cost of disagreement. It is therefore interesting to consider situations in which players can influence the extent of the social cost of disagreement in order to sway the bargaining outcome in their favor.

One classical economic example is the threat by a trade union to initiate a strike unless its demands are met in a wage negotiation (see, for instance, Fernandez and Glazer (1991)). In international relations, one example could be a country that threatens to walk away from the bargaining table for some time if certain red lines are crossed (Li (2011)). The bargaining literature has long since recognized that such commitments, postures, and threats are highly effective tactics at the negotiating table (see Schelling (1956, 1960), Muthoo (1992, 1996), Abreu and Gul (2000), and Ellingsen and Miettinen (2008)). Much of the analysis in the literature focuses on the credibility of commitments and threats. What is oftentimes neglected, however, is that credible threats may be harmful even if they are not executed. To be more precise, suppose that the surplus consists of the gains from some kind of cooperation between the players. These gains may crucially depend on good relations of trust among the parties, and may thus be compromised by aggressive and threatening behavior.

To be more specific, consider the example of a labor dispute. The firm might have the ability to bargain aggressively and obtain agreement on a very low wage. However, it needs to take into account that its success also depends on many non–contractible forms of cooperation or goodwill by its employees, and could therefore be hurt when employees feel treated unfairly.

Similarly, a union could decide to bargain aggressively and threaten a strike. Even if an agreement is reached and the strike never actually takes place, the threat might have scared away customers and thus done damage to the firm's profitability.

As a final example, aggressive bargaining behavior in international negotiations may damage prospects for future cooperation between the countries involved, and thus be detrimental to the "surplus" that parties can generate. Moreover, when countries threaten each other with sanctions or war, the mere threat may scare away investors and harm economic development. In a nutshell, many such bargaining situations involve welfare considerations while the literature has often treated them as mere surplus division problems.

Our aim in the present paper is to model a unanimity bargaining game which can capture trade–offs of the type illustrated above: Proposers are able to threaten other players with a high social cost of disagreement. The more a proposer exploits this power, however, the smaller is the available surplus. More specifically, we consider the following model: When making a proposal, the proposer can endogenously choose the time lapse which occurs in case this proposal is rejected. Due to players' impatience, choosing a long time lapse is akin to making a harsh threat. The size of the available surplus decreases with the length of the chosen time lapse. Thus, the proposer can strengthen her individual bargaining position by diminishing the surplus. This behavior is a form of "rent–seeking."

Our bargaining model has the following additional features: We allow for any finite number of players. They share a common rate of time preference. A player who rejects a proposal has the right to make a counter–proposal. This rejector–proposes bargaining protocol is appealing for several reasons: It is a proper generalization of Rubinstein's alternating offers protocol to the case with more than two players. Moreover, it treats all responding players symmetrically. Following the work of Selten (1981), the rejector–proposes protocol has been commonly used in a variety of bargaining games involving both surplus division and coalition formation problems, examples include Chatterjee et al. (1993), Bloch (1996), Ray and Vohra (1999), Ray (2007), and Kawamori (2013). In addition, Kawamori (2008) and Britz et al. (2014) study a class of bargaining protocols that generalizes the rejector–proposes protocol. In accordance with the standard approach in the literature, we restrict attention to subgame–perfect equilibria in stationary strategies.

More specifically, we focus on two main questions: We investigate the equilibrium level of surplus destruction, and we assess the extent of the endogenously determined proposer premium, that is, the extra payoff that the proposer can obtain compared to other players.

Our main results are as follows: There is a unique equilibrium prediction for the level of surplus destruction as well as for the shares of surplus allocated to each player. Agreement is always immediate. If players are either very patient or very impatient, no surplus destruction occurs so that equilibrium is efficient. If players' rate of time preference falls within an intermediate range, some surplus destruction does occur in equilibrium. Within this range, there is a unique value for the rate of time preference at which surplus destruction peaks.

The intuition underlying this non–monotonicity is as follows: If players are very patient, then the threat of a lengthy delay is not very effective, and it is not worthwhile to waste surplus for it. On the other hand, if players are very impatient, then the prospect of delay is a very effective threat, even if the prospective delay is only quite short. Hence, it does not make sense to waste a lot of surplus in order to create the threat of a very lengthy delay.

We establish a tight upper bound on the share of the surplus which is destroyed in equilibrium. Although equilibrium surplus destruction depends non–monotonically on the rate of time preference, it is still true that the proposer premium is monotonically increasing in the rate of time preference. Moreover, we demonstrate that, in the presence of equilibrium surplus destruction, the proposer receives a greater payoff than in standard Rubinstein bargaining.

In the bargaining literature, several papers have considered the possibility that a proposer can impose a social cost after the rejection of his proposal. Haller and Holden (1990), Fernandez and Glazer (1991), and Manzini (1999) model wage bargaining between a union and a firm where the firm can go on strike after a disagreement. Avery and Zemsky (1994) and Busch et al. (1998) are interested in the possibility of "burning money" after the rejection of a proposal. Li (2011) allows a proposer to suspend bargaining for some time after an unsuccessful round. One important feature of these earlier papers is that, prior to the rejection of the current proposal, the proposer cannot commit to imposing a cost in case of rejection. Therefore, the main question in a model like Li's is how the proposer can give credibility to the threat of suspending the negotiation. Given that Li (2011) analyzes subgame–perfect equilibria in a setting with two players, this credibility comes from a cascade of punishment modes that the game enters after deviations from the equilibrium strategies. The approach in the present paper is different: We are not concerned with the credibility of the threat to delay. Instead, we assume that the proposer simultaneously chooses the proposal and the length of the time lapse which occurs in case of rejection. Due to this assumption, the threat of delay is perfectly credible. However, it comes at the cost of diminishing the surplus. This cost limits the leverage that the proposer can gain by such a threat. Another difference between Li (2011) and the present paper is that we allow for an arbitrary number of players, and therefore need to restrict ourselves to equilibria in stationary strategies. Hence, punishment modes after deviations as in Li (2011) are not relevant in our context.

In this way, our work is related to the idea of costly commitments. Bargaining games in which players incur a private cost to credibly commit themselves have been studied repeatedly in the literature. For instance, Cunyat (2004) considers the possibility that players make pre–bargaining commitments. They choose the probability with which their commitments are binding, and they incur a cost which is proportional to the chosen probability. Ellingsen and Miettinen (2008) contribute to the debate on inefficient delays in bargaining using a model with costly commitments. In Miettinen and Perea (2015), players choose their commitments in each bargaining round, and they pay a cost which is proportional to the amount of surplus that they are committing to.

In another related stream of literature, Evans (1997), Yildirim (2007, 2010), Imai and Salonen (2012), and Cardona and Polanski (2013) study bargaining games in which players engage in a contest in order to become the proposer. The player who exerts most costly effort in this contest has the highest probability of becoming the proposer. One interpretation of this model is that bargaining happens through a mediator, and players use "lobbying" to compete for the favorable attention of that mediator (Yildirim (2007, 2010)). Our work complements this approach: We adopt from these papers the idea that players' bargaining power is endogenously determined through a kind of rent–seeking behavior. We differ from the aforementioned work in two important respects: First, what we make endogenous is not proposer recognition but the strength of the proposer's bargaining position. Second, the rent–seeking in our model does not come at a private cost to the proposer, but diminishes the surplus.[1]

The remainder of the paper is organized as follows: Section 2 contains the formal description of the bargaining game. In Section 3, we provide a characterization of subgame–perfect equilibria in stationary strategies. These equilibria may be efficient, or may involve surplus destruction. These two cases are discussed in Sections 4 and 5, respectively. In Section 6, we focus on the relationship between basic model parameters and the extent of equilibrium surplus destruction. Section 7 concludes.

3.2 Model description

We consider a unanimity bargaining game with players $N = \{1, \ldots, n\}$ who are negotiating on the division of a surplus. Bargaining proceeds in *rounds.* The first bargaining round takes place at time $\tau = 0$. We will describe shortly how the timing of any further bargaining rounds is determined endogenously. The size of the surplus at any time $\tau \geq 0$ is denoted by Π^τ. The initial surplus size Π^0 is normalized to one, while the size Π^τ of the surplus at any time $\tau > 0$ is endogenously determined.

Consider a bargaining round which takes place at time $\tau \geq 0$, when the current surplus is of size Π^τ. Any such bargaining round has the following structure: One player is the proposer, let us say for the moment that it is Player i. The proposer ("she") chooses a pair (θ^i, t^i), where the proposal $\theta^i = (\theta^i_1, \ldots, \theta^i_n)$ specifies the share of the current surplus Π^τ which Player i offers to each Player $1, \ldots, n$. The proposer's choice of t^i has two effects: First, it determines the share $1 - \sigma t^i$ of the surplus that can actually be divided, while the share σt^i is destroyed, where $\sigma > 0$. Second, it implies that a rejection of the proposal θ^i leads to a time lapse of length $\Delta + t^i$ before the next bargaining round, where $\Delta > 0$. Intuitively, the proposer threatens other players with a delay of $\Delta + t^i$ in case of a rejection. In order to make this threat, he sacrifices the share σt^i of the surplus. The parameter $\sigma > 0$ determines how detrimental the proposer's threat is for the surplus, while the parameter $\Delta > 0$ represents the minimal increment of time which has to elapse between two bargaining rounds, regardless of the proposer's actions.[2] We require the

[1] In Appendix D, we briefly consider a variant of our model in which this assumption is reversed, and the proposer privately bears the cost of rent–seeking. We argue that important qualitative results of our analysis carry over to that setting. This version of our model could be given the following interpretation similar to that in Yildirim (2007, 2010): Bargaining takes place through a mediator. The proposer is the player who currently has the mediator's ear. The proposer's objective is to intensify the relationship with the mediator through lobbying, which involves costly effort.

[2] Indeed, we assume that even in the absence of any surplus destruction by the proposer, it takes some

pair (θ^i, t^i) chosen by the proposer to be *feasible.* This means that $t^i \in [0, 1/\sigma]$, the proposal θ^i is non–negative in all components, and $\sum_{j \in N} \theta^i_j \leq 1 - \sigma t^i$. In principle, it is feasible for Player i to destroy the whole surplus. That is, she might choose $t^i = 1/\sigma$ and $\theta^i = (0, \ldots, 0)$. In that case, the game ends immediately, and all players receive zero payoffs. If Player i has chosen some feasible pair (θ^i, t^i) such that $t^i < 1/\sigma$, then the bargaining round continues as follows: Players $1, \ldots, n$ respond sequentially to the proposal θ^i by *acceptance* or *rejection.*[3] If all players accept θ^i, then the game ends, and each Player $j \in N$ receives the share θ^i_j of the current surplus Π^τ. As soon as some Player $j \in N$ rejects θ^i, the current bargaining round ends, and Player j becomes the proposer of the next bargaining round which takes place at time $\tau + \Delta + t^i$. The remaining surplus in the next bargaining round is $\Pi^{\tau+\Delta+t^i} = (1 - \sigma t^i)\Pi^\tau$.

All players are risk–neutral and impatient. They share a common rate of time preference $r > 0$. Thus, if Player j receives a share θ^i_j of the surplus Π^τ in a bargaining round at time τ, then this corresponds to a payoff of $e^{-r\tau}\theta^i_j\Pi^\tau$ for Player j. If no proposal is ever accepted ("perpetual disagreement"), then all players receive zero payoffs.

In what follows, we denote by $G^i(\Delta, n, r, \sigma)$ a subgame of our bargaining game which starts at a history where Player i is the proposer and chooses (θ^i, t^i). Note that all such subgames are equivalent up to the absolute size of the surplus at their roots. We have specified that whenever Player i rejects a proposal, then a subgame $G^i(\Delta, n, r, \sigma)$ follows after some time lapse. This "rejector proposes" bargaining protocol is a proper generalization of Rubinstein's alternating offers bargaining to n players. In order to complete the description of the entire game, we only have left to specify the identity of the initial proposer. Without loss of generality, let us say that Player 1 proposes first, so that the entire game $G(\Delta, n, r, \sigma)$ is a subgame of the type $G^1(\Delta, n, r, \sigma)$.

Let us now turn to the equilibrium concept. It is well–known that an analysis of subgame–perfect equilibrium does not lead to sharp payoff predictions in unanimity bargaining games with at least three players. Herrero (1985) and Haller (1986) have provided evidence for a "folk theorem" in multilateral bargaining: Every individually rational payoff allocation is supported by some subgame–perfect equilibrium.[4] Aiming for sharp payoff predictions, the bargaining literature has recognized subgame–perfect equilibrium

small increment $\Delta > 0$ of time before any other player can make a new proposal. In the sequel of the paper, we will pay special attention to the case where Δ is sufficiently small. This is consistent with the approach in the literature on unanimity bargaining games in the tradition of Rubinstein (1982). The analysis of that class of games relies on the presence of at least some small exogenously given cost of delay ("bargaining friction"), and often focuses on the case where this cost tends to zero.

[3]The order in which players respond to a proposal is not crucial for the results. In order to express the optimal accept/reject decisions in a straightforward manner, it is convenient to assume that players respond in ascending order $1, 2, \ldots, n$ regardless of the current proposer's identity.

[4]More recently, Herings et al. (2017) have established this folk theorem for a general class of unanimity bargaining games with sequential accept/reject decisions, provided that the discount factor is close enough to one. Their proof is constructive, and only makes use of strategies with one–period recall. A slight modification of the construction in Herings et al. could be used to establish an analogous folk theorem in our model, provided that r is small enough.

in stationary strategies as the appropriate solution concept for multilateral unanimity bargaining games. Stationarity requires a player to make the same proposal whenever she is the proposer, and to condition accept/reject decisions only on the current proposal, rather than on any other aspects of the history of play. In the game at hand, the endogenous choice of the time lapse between bargaining rounds requires a modified notion of a stationary strategy. In particular, we allow a player to condition an accept/reject decision on the current proposal and on the time lapse chosen by the current proposer. In turn, stationarity requires that a player choose the same proposal and the same time lapse whenever she is the proposer.

More formally, a *stationary strategy* for a Player $i \in N$ consists of a feasible pair (θ^i, t^i) which Player i chooses whenever she is the proposer, and of a correspondence $A^i(t)$ such that Player i responds to proposer j's choice of (θ^j, t^j) by acceptance if $\theta^j \in A^i(t^j)$, and by rejection otherwise. Henceforth, we denote a stationary strategy of Player i concisely by $(\theta^i, t^i, A^i(t))$.[5] A *stationary subgame–perfect Nash equilibrium (SSPE)* is a profile of stationary strategies which is a subgame–perfect Nash equilibrium.

A profile of stationary strategies $(\theta^i, t^i, A^i(t))_{i\in N}$ induces a set $A(t) = \bigcap_{i\in N} A^i(t)$ of unanimously acceptable proposals for each time lapse t. It is crucial that a stationary strategy specifies the proposal, the surplus destruction, and the accept/reject decisions only in terms of shares of the current[6] surplus, independently of its absolute size. This allows us to define a quantity $\gamma_k((\theta^i, t^i, A^i(t))_{i\in N})$ as follows: Consider any subgame $G^k(\Delta, n, r, \sigma)$, and assume that the appropriate restriction of the profile $(\theta^i, t^i, A^i(t))_{i\in N}$ of stationary strategies is played in this subgame. Then, $\gamma_k((\theta^i, t^i, A^i(t))_{i\in N})$ is the resulting payoff of Player k, expressed as a share of the surplus at that subgame's root. The quantity $\gamma_k((\theta^i, t^i, A^i(t))_{i\in N})$ is crucial to define the cut–off below which it is optimal for Player k to reject a proposal. Indeed, suppose that the current proposer has chosen some $\widehat{t} \in [0, 1/\sigma]$. Suppose furthermore that after a rejection of the current proposal, the profile $(\theta^i, t^i, A^i(t))_{i\in N}$ will be played. Then, the continuation payoff to Player k from rejecting the current proposal corresponds to a share $(\ 1 - \sigma\widehat{t}\ )e^{-r(\ \Delta+\widehat{t}\ )}\gamma_k((\theta^i, t^i, A^i(t))_{i\in N})$ of the current surplus. The payoff from rejecting and becoming the proposer in the next round depends on the current proposer's choice of the time lapse in two ways: First, the factor $e^{-r(\ \Delta+\widehat{t}\ )}$ is applied because of the delay occurring after a rejection of the current proposal. It reflects actual discounting of future payoffs by impatient players. Second, the factor $(\ 1 - \sigma\widehat{t}\ )$ is applied because any payoffs in the next bargaining round are expressed as shares of a surplus which has shrunk by the share $\sigma\widehat{t}$ due to the current proposer's choice of $\widehat{t}$. We can interpret the entire term $(\ 1 - \sigma\widehat{t}\ )e^{-r(\ \Delta+\widehat{t}\ )}$ as an "implicit discount factor,"

[5]The continuation game after Player i's rejection of a proposal is independent of the current proposer's identity. Thus, it is appropriate to define a stationary strategy such that the player cannot condition his acceptance or rejection on the identity of the current proposer.

[6]Throughout the paper, we mean by the "current surplus" the surplus at the beginning of the current round.

arising endogenously from the proposer's choice of the time lapse. Loosely speaking, one could say that the proposer is effectively choosing the other players' discount factor. One may wonder whether the model could equivalently be rephrased in such a way that the proposer can change the responding players' rate of time preference while the time lapse between rounds is fixed at Δ. While such a reformulation would be feasible from a technical point of view, it is at odds with the idea that the players' (time) preferences are part of the primitives of the model.

3.3 A Characterization of SSPE

This section is devoted to the characterization of SSPE, which is the basis for the analysis of equilibrium surplus destruction in the sequel of the paper. In particular, we show the following theorem:

Theorem 3.1. *The profile $(\theta^i, t^i, A^i(t))_{i\in N}$ is an SSPE if and only if it satisfies the following set of conditions for all $i \in N$:*

$$\begin{aligned} \theta^i_i &= 1 - \sigma t^i - \sum_{j\in N\setminus\{i\}} \theta^i_j, & (3.1)\\ \theta^i_j &= (1-\sigma t^i)e^{-r(\Delta+t^i)}\theta^j_j,\ \forall j \in N\setminus\{i\}, & (3.2)\\ A^i(\,\widehat{t}\,) &= \{v \in \mathbb{R}^+_n | v_j \geq (1-\sigma\widehat{t}\,)e^{-r(\Delta+\widehat{t}\,)}\theta^j_j,\ \forall j \in \{i,\dots,n\}\},\ \forall\,\widehat{t} \in [0, 1/\sigma], & (3.3)\\ t^i &\in \arg\max_{t\in[0,1/\sigma]} (1-\sigma t)\left(1 - e^{-r(\Delta+t)}\sum_{j\in N\setminus\{i\}} \theta^j_j\right). & (3.4)\end{aligned}$$

Moreover, on the path of play of any SSPE, agreement is reached immediately in every subgame.

The equilibrium conditions (3.1)–(3.3) have interpretations which are familiar from the literature on unanimity bargaining games. In particular, Eqn. (3.1) says that the part of the surplus which is not destroyed is completely distributed to the players. Eqn. (3.2) indicates that each player other than the proposer is offered the share which makes him exactly indifferent between acceptance and rejection. Eqn. (3.3) says that a player accepts a proposal if and only if each player who has not yet accepted is offered at least his continuation payoff from becoming the proposer in the next round.[7]

The proof of Theorem 3.1 is decomposed into several parts. In Appendix A, we establish a sequence of auxiliary results which demonstrate the following two statements: First, Eqns. (3.1)–(3.3) are necessary conditions for an SSPE. Second, in an SSPE, the immediate agreement property holds in every subgame. In what follows, we focus on two

[7] The following intuition is useful to understand the acceptance sets: If Player i anticipates that some Player $i+1, i+2, \dots$ would reject the current proposal, then Player i is better off rejecting it himself. This is formally shown in Appendix A.

additional claims: We argue that (3.4) is another necessary condition for SSPE, and we point out that the necessary conditions (3.1)–(3.4) for an SSPE are also sufficient. These claims combined amount to the proof of Theorem 3.1.

Indeed, let us now focus on condition (3.4) above. When Player i is the proposer, she faces the following trade–off: If she chooses t^i, then she must offer the other players the share $(1-\sigma t^i)e^{-r(\Delta+t^i)}\sum_{j\in N\setminus\{i\}}\theta_j^j$ in order to obtain their agreement. This term is clearly decreasing in t^i. Choosing a greater time lapse improves Player i's bargaining position relative to the other players. On the downside, however, choosing t^i implies that a share σt^i of the surplus is destroyed. More formally, given proposals θ^j of all other players $j \in N \setminus \{i\}$, and given that each of these proposals is unanimously accepted, Player i's payoff can be thought of as the following function of t^i:

$$\xi^i(t^i) = (1-\sigma t^i)\left(1 - e^{-r(\Delta+t^i)}\sum_{j\in N\setminus\{i\}}\theta_j^j\right).$$

Hence, we see that condition (3.4) above is necessary for an SSPE, as desired.

Finally, we argue that the necessary conditions (3.1)–(3.4) for an SSPE are also sufficient. Eqn. (3.4) ensures that a proposer has no profitable deviation involving a different choice of t^i. Eqns. (3.1)–(3.2) imply that a proposer cannot profitably deviate by changing only θ^i, but not t^i. We have left to verify that accept/reject decisions are optimal in a profile of stationary strategies which satisfies conditions (3.1)–(3.4). This is shown in Lemma 3.12 in Appendix A, and completes the proof of Theorem 3.1.

Note that setting all the time lapses $t^1,\ldots,t^n$ in Eqns. (3.1)–(3.2) equal to zero yields the equations which are familiar from the equilibrium analysis of standard unanimity bargaining games with an exogenously fixed time lapse $\Delta > 0$ and concomitant discount factor $e^{-r\Delta} < 1$. The novel element of the above equilibrium characterization is the optimization problem (3.4) which effectively endogenizes the discount factor. We will see that two kinds of solution to this optimization problem are relevant for the analysis: First, a corner solution of the optimization problem corresponds to the case where $t^i = 0$ is chosen. Second, an interior solution corresponds to the case where some $t^i > 0$ is chosen.

More formally, consider the optimization problem (3.4). Take the derivatives

$$\begin{aligned}
\partial\xi^i(t)/\partial t &= -\sigma + (\sigma + r(1-\sigma t))\, e^{-r(\Delta+t)}\left(\sum_{j\in N\setminus\{i\}}\theta_j^j\right)\\
\partial^2\xi^i(t)/\partial^2 t &= -r\sigma e^{-r(\Delta+t)}\left(\sum_{j\in N\setminus\{i\}}\theta_j^j\right) - r(\sigma + r(1-\sigma t))e^{-r(\Delta+t)}\left(\sum_{j\in N\setminus\{i\}}\theta_j^j\right).
\end{aligned}$$

We see that

$$\partial^2\xi^i(t)/\partial^2 t < 0 \text{ for } t \in [0, 1/\sigma].$$

Moreover, evaluating the first–order derivative at the points $t = 0$ and $t = 1/\sigma$, we see that

$$\begin{aligned} \partial\xi^i(t)/\partial t_{|_{t=0}} &= -\sigma + (\sigma + r)e^{-r\Delta}\left(\sum_{j\in N\setminus\{i\}} \theta^j_j\right), \\ \partial\xi^i(t)/\partial t_{|_{t=1/\sigma}} &= -\sigma + \sigma e^{-r(\Delta+1/\sigma)}\left(\sum_{j\in N\setminus\{i\}} \theta^j_j\right). \end{aligned}$$

We can conclude that there are two possible cases:

1. If $e^{-r\Delta}\left(\sum_{j\in N\setminus\{i\}} \theta^j_j\right) \leq \sigma/(\sigma + r)$, then the first–order derivative $\partial\xi^i(t)/\partial t$ is non–positive for $t \in [0, 1/\sigma]$ and strictly negative for $t \in (0, 1/\sigma)$. Hence, the optimization problem (3.4) has only a corner solution at $t = 0$.

2. If $e^{-r\Delta}\left(\sum_{j\in N\setminus\{i\}} \theta^j_j\right) > \sigma/(\sigma+r)$, then the first–order derivative $\partial\xi^i(t)/\partial t$ is strictly positive at $t = 0$, and strictly monotonically decreasing on the interval $[0, 1/\sigma]$. If $\partial\xi^i(t)/\partial t_{|_{t=1/\sigma}} \geq 0$, then we would find a corner solution at $t = 1/\sigma$. However, we have shown in Corollary 3.3 (in Appendix A) that this does not occur in an SSPE. So we need to consider only the case where $\partial\xi^i(t)/\partial t_{|_{t=1/\sigma}} < 0$, in which we find a unique interior solution at some $t \in (0, 1/\sigma)$. This solution is given by the first–order condition $\partial\xi^i/\partial t = 0$, which can be written as $\sum_{j\in N\setminus\{i\}} \theta^j_j = \frac{\sigma e^{r(\Delta+t)}}{\sigma+r(1-\sigma t)}$.

We have found that Player i chooses the level of surplus destruction and the concomitant time lapse depending on the basic model parameters and on the value of $\left(\sum_{j\in N\setminus\{i\}} \theta^j_j\right)$. Since each responding player becomes the next proposer if he rejects the current proposal, this quantity is the (un–discounted) sum of the reservation payoffs of all players. It can be interpreted as the "price of agreement."

Theorem 3.2. *Let $(\theta^k, t^k, A^k(t))_{k\in N}$ be an SSPE. It holds that either $\sum_{j\in N\setminus\{i\}} \theta^j_j \leq \frac{\sigma e^{r\Delta}}{\sigma+r}$ and $t^i = 0$, or $\sum_{j\in N\setminus\{i\}} \theta^j_j > \frac{\sigma e^{r\Delta}}{\sigma+r}$ and $t^i > 0$. In the latter case, $t^i > 0$ is the unique solution to $\sum_{j\in N\setminus\{i\}} \theta^j_j = \frac{\sigma e^{r(\Delta+t)}}{\sigma+r(1-\sigma t)}$.*

In the sequel of the paper, we only deal with profiles of stationary strategies in which the acceptance sets are as specified in Eqn. (3.3). Therefore, it will be convenient to ease the notation and refer to such profiles of stationary strategies only as $(\theta^i, t^i)_{i\in N}$.

In the next section, we are going to discuss SSPE in which all players choose zero surplus destruction. We call such equilibria *efficient SSPE.* In Section 5, we then turn to SSPE in which at least one player chooses to destroy some surplus. We call such equilibria *SSPE with surplus destruction.* In the sequel of the paper, we show that all SSPE are "symmetric," that is, all players choose the same surplus destruction. Moreover, we show that for any given choice of the model parameters, there exists either an efficient SSPE, or an SSPE with surplus destruction, but not both.

3.4 Efficient SSPE

In this section, we study SSPE in which all players choose not to destroy any surplus. The following theorem specializes the characterization of an SSPE to this case.

Theorem 3.3. *The proposals $\theta^1, \ldots, \theta^n$ are part of an efficient SSPE if and only if the following conditions hold for all $i \in N$ and all $j \in N \setminus \{i\}$:*

$$\theta_i^i = 1 - \sum_{j \in N \setminus \{i\}} \theta_j^i, \tag{3.5}$$

$$\theta_j^i = e^{-r\Delta} \theta_j^j, \tag{3.6}$$

$$\theta_i^i \geq r/(\sigma + r). \tag{3.7}$$

Proof. If $t^i = 0$, then Eqns. (3.1)–(3.2) specialize to Eqns. (3.5)–(3.6). Moreover, due to Theorem 3.2, the inequality $e^{-r\Delta}(\sum_{j \in N \setminus \{i\}} \theta_j^j) \leq \sigma/(\sigma + r)$ is satisfied in an efficient SSPE. Combined with Eqns. (3.5)–(3.6), this implies Ineq. (3.7) above. □

Eqns. (3.5)–(3.6) above amount to a system of n^2 independent equations which allow us to solve uniquely for proposals $(\theta^1, \ldots, \theta^n)$. This solution is

$$\theta_i^i = \frac{1}{1 + (n-1)e^{-r\Delta}}, \ \forall i \in N,$$

$$\theta_j^i = \frac{e^{-r\Delta}}{1 + (n-1)e^{-r\Delta}}, \ \forall i \in N, \ \forall j \in N \setminus \{i\}.$$

In the sequel of the paper, we will sometimes compare the bargaining game under consideration to the following benchmark: The *bargaining game with fixed timing* is like the bargaining game in the present paper, except that all proposers are constrained not to destroy surplus, and thus to choose $t^1 = \ldots = t^n = 0$. The time lapse between bargaining rounds of a bargaining game with fixed timing is of length Δ, and the concomitant discount factor is $e^{-r\Delta}$. The bargaining game with fixed timing is a special case of the game analyzed in Britz et al. (2014). It follows from the analysis in that paper that, in the stationary equilibrium of the bargaining game with fixed timing, every proposer offers $\frac{1}{1+(n-1)e^{-r\Delta}}$ to herself and $\frac{e^{-r\Delta}}{1+(n-1)e^{-r\Delta}}$ to every responding player, and agreement is reached immediately in every subgame. Henceforth, we refer to the allocation which gives the proposer a share $\frac{1}{1+(n-1)e^{-r\Delta}}$ of the surplus, and which gives each player other than the proposer a share $\frac{e^{-r\Delta}}{1+(n-1)e^{-r\Delta}}$ of the surplus as the *fixed timing allocation.* This allocation will be an important benchmark in our later analysis of SSPE with surplus destruction.

Theorem 3.3 has the following implication: Any efficient SSPE leads to the fixed timing allocation. However, depending on the basic model parameters, the fixed timing allocation may or may not be supported by an SSPE. More specifically, for the fixed timing allocation to be supported by an SSPE, an additional restriction given by Ineq. (3.7) must be satisfied. This inequality imposes a lower bound on the proposer's share, given the parameters r and σ. The intuition is as follows: If the fixed timing allocation does not give the proposer a sufficiently large share of the surplus, then the proposer has an incentive to choose a strictly positive level of surplus destruction in order to enhance her bargaining power, and thus depart from the fixed timing allocation.

Next, we state a necessary and sufficient condition for the existence of an efficient SSPE.

Lemma 3.1. *An efficient SSPE exists if and only if $e^{r\Delta}/r \geq (n-1)/\sigma$.*

Proof. *If.* Consider proposals $(\bar{\theta}^i)_{i\in N}$ defined for every $i \in N$ by $\bar{\theta}^i_i = \frac{1}{1+(n-1)e^{-r\Delta}}$ and $\bar{\theta}^i_j = \frac{e^{-r\Delta}}{1+(n-1)e^{-r\Delta}}$ for $j \in N \setminus \{i\}$. These proposals satisfy Eqns. (3.5)–(3.6). If $e^{r\Delta} \geq (n-1)(r/\sigma)$, then $e^{-r\Delta}(n-1) \leq \sigma/r$, and so $\bar{\theta}^i_i \geq \frac{1}{1+\sigma/r} = \frac{r}{r+\sigma}$, and so Ineq. (3.7) is satisfied. Indeed, we have shown that proposals $(\bar{\theta}^i)_{i\in N}$ are part of an efficient SSPE.

Only If. Suppose that $(\bar{\theta}^i)_{i\in N}$ are proposals of an efficient SSPE. Then, for any $i \in N$, we have $1 = \theta^i_i + \sum_{j\in N\setminus\{i\}} \theta^i_j = \theta^i_i + e^{-r\Delta} \sum_{j\in N\setminus\{i\}} \theta^j_j$. Due to Ineq. (3.7), it holds that $\theta^k_k \geq r/(\sigma+r)$ for every $k \in N$, so we have the inequality $\left(\frac{r}{\sigma+r}\right) + (n-1)e^{-r\Delta}\left(\frac{r}{\sigma+r}\right) \leq 1$. Equivalent transformation yields $e^{r\Delta}/r \geq (n-1)/\sigma$, as desired.

□

Note that the term $e^{r\Delta}/r$ grows without bound in the limit as r goes to zero as well as in the limit as r goes to infinity. Thus, an efficient SSPE exists for sufficiently small as well as for sufficiently large r. The intuition is as follows: The proposer derives bargaining power from the time lapse which would occur if her proposal was rejected. The greater is the rate of time preference, the more bargaining power the proposer can derive from this time lapse. If r is very large, then even the minimal time lapse $\Delta > 0$ confers so much bargaining power to the proposer that she does not find it worthwhile to destroy surplus in order to gain additional bargaining power. If r is very small relative to σ, then she could only improve her bargaining position by destroying an unduly large amount of surplus.

One implication of the analysis in this section is that if Δ is too large relative to other model parameters, then an efficient SSPE exists, and we are back to the analysis of a bargaining game with fixed timing. However, we are going to demonstrate in the sequel of the paper that, for small values of Δ, there is an intermediate range of r so that surplus destruction does occur in equilibrium. Within such an intermediate range, the following two conditions hold: On the one hand, r is small enough so that the minimal time lapse Δ by itself does not generate much bargaining power for the proposer. On the other hand, r is large enough so that the amount of surplus destruction needed to achieve some increase in bargaining power is not prohibitive. Due to the surplus destruction, our results are going to depart from those pertaining to bargaining games with fixed timing.

3.5 SSPE with Surplus Destruction

As a next step, we are going to complement the previous section with an analysis of SSPE with surplus destruction. By definition, an SSPE with surplus destruction is an SSPE in which at least one player destroys surplus. One of the results in this section is that all SSPE are "symmetric," that is, in an SSPE, all players choose the same surplus destruction and thus the same time lapse. We are going to show that, for any choice of the basic model parameters Δ, n, σ, and r, there exists a unique SSPE. We establish conditions on the basic model parameters under which this SSPE is efficient or involves surplus destruction. In the latter case, we derive an equality which implicitly determines the amount of surplus destruction.

As in the previous section, we begin the analysis with a statement which specializes Theorem 3.1.

Lemma 3.2. *Suppose that $(\theta^k, t^k)_{k \in N}$ is an SSPE with surplus destruction. Then, the following equalities hold for every $i \in N$ such that $t^i > 0$:*

$$\theta_i^i = \frac{r(1-\sigma t^i)^2}{\sigma + r(1-\sigma t^i)}, \tag{3.8}$$

$$\sum_{j \in N \setminus \{i\}} \theta_j^i = \frac{\sigma(1-\sigma t^i)}{\sigma + r(1-\sigma t^i)}, \tag{3.9}$$

$$\sum_{j \in N \setminus \{i\}} \theta_j^j = \left(\frac{\sigma}{\sigma + r(1-\sigma t^i)}\right) e^{r(\Delta + t^i)}. \tag{3.10}$$

Proof. If $t^i > 0$ in an SSPE, then $t^i > 0$ must be an interior solution to the optimization problem (3.4). The corresponding first–order condition is Eqn. (3.10). From Eqns. (3.1)–(3.2), we find $\theta_i^i = 1 - \sigma t^i - (1-\sigma t^i)e^{-r(\Delta+t^i)} \sum_{j \in N\setminus\{i\}} \theta_j^j$. By substitution from Eqn. (3.10) above, we find

$$\theta_i^i = 1 - \sigma t^i - \frac{\sigma(1-\sigma t^i)}{\sigma + r(1-\sigma t^i)} = \frac{r(1-\sigma t^i)^2}{\sigma + r(1-\sigma t^i)},$$

thus establishing Eqn. (3.8). Combining the above equality with Eqn. (3.1), we find Eqn. (3.9), as desired. □

The next lemma claims that SSPE are "symmetric," that is, all players choose the same level of surplus destruction.

Lemma 3.3. *If* $(\theta^k, t^k)_{k\in N}$ *is an SSPE, then* $t^1 = t^2 = \ldots = t^n$.

The proof of Lemma 3.3 is provided in Appendix B.

Consider an SSPE $(\theta^k, t^k)_{k\in N}$ with surplus destruction. We have shown that all players choose the same strictly positive time lapse.

Using Eqns. (3.2) and (3.8), we can conclude that if $(\theta^k, t^k)_{k\in N}$ is an SSPE with surplus destruction, then there is a "proposer share" $x > 0$ such that $x = \theta^i_i$ for every $i \in N$, and there is a "responder share" $y > 0$ such that $y = \theta^i_j$ for every $i \in N$ and $j \in N \setminus \{i\}$. From Eqns. (3.8)–(3.10) above, we can infer that in any SSPE with surplus destruction, the proposer share x, the responder share y, and the time lapse t chosen by each player satisfy the following equalities:

$$x/y = e^{r(\Delta+t)}(1-\sigma t)^{-1}, \tag{3.11}$$

$$x/y = \left(\frac{r}{\sigma}\right)(1-\sigma t)(n-1), \tag{3.12}$$

and, consequently, the equilibrium choice of the time lapse t satisfies

$$e^{r(\Delta+t)} = \left(\frac{r}{\sigma}\right)(1-\sigma t)^2(n-1).$$

Now we can state our results concisely as the following characterization of SSPE with surplus destruction:

Theorem 3.4. *In an SSPE with surplus destruction, every player chooses the time lapse* t *such that*

$$e^{r(\Delta+t)} = \left(\frac{r}{\sigma}\right)(1-\sigma t)^2(n-1). \tag{3.13}$$

Moreover, every Player $i \in N$ *chooses the proposal* θ^i *given by:*

$$\theta^i_i = \frac{r(1-\sigma t)^2}{\sigma + r(1-\sigma t)}, \tag{3.14}$$

$$\theta^i_j = \left(\frac{\sigma(1-\sigma t)}{\sigma + r(1-\sigma t)}\right) / \,(n-1)\,, \ \forall j \in N \setminus \{i\}. \tag{3.15}$$

Consider Eqn. (3.13) which determines the time lapse in an SSPE with surplus destruction. The left–hand side of the equality is strictly monotonically increasing on $t \in [0, 1/\sigma]$, and equals $e^{r\Delta}/r$ at the point $t = 0$. The right–hand side is strictly monotonically decreasing on $t \in [0, 1/\sigma]$, and equals $\left(\frac{r}{\sigma}\right)(n-1)$ at the point $t = 0$. For surplus destruction to occur in an SSPE, some $t > 0$ must solve Eqn. (3.13). Hence, we have the following necessary and sufficient condition for the existence of an SSPE with surplus destruction:

Lemma 3.4. *An SSPE with surplus destruction exists if and only if $e^{r\Delta}/r < (n-1)/\sigma$.*

Proof. *If.* Suppose that $e^{r\Delta}/r < (n-1)/\sigma$. Then, there is $\bar{t} > 0$ such that $e^{r(\Delta+\bar{t})} = \left(\frac{r}{\sigma}\right)(1-\sigma\bar{t})^2(n-1)$. For every $i \in N$, define $\bar{\theta}^i_i = \frac{r(1-\sigma\bar{t})^2}{\sigma+r(1-\sigma\bar{t})}$, and $\bar{\theta}^i_j = \left(\frac{\sigma(1-\sigma\bar{t})}{\sigma+r(1-\sigma\bar{t})}\right)\left(\frac{1}{n-1}\right)$ for $j \in N \setminus \{i\}$. Moreover, let $\bar{t}^k = \bar{t}$ for every $k \in N$. It is now easily verified that $(\bar{\theta}^k, \bar{t}^k)_{k\in N}$ is an SSPE.

Only If. Suppose that $(\bar{\theta}^k, \bar{t}^k)_{k\in N}$ is an SSPE. Then, for every $k \in N$, the equality $e^{r(\Delta+\bar{t}^k)} = \left(\frac{r}{\sigma}\right)(1-\sigma\bar{t}^k)^2(n-1)$ is satisfied. Since $\bar{t}^k \in (0, 1/\sigma]$, it holds that $\left(\frac{r}{\sigma}\right)(1-\sigma\bar{t}^k)^2(n-1) < \left(\frac{r}{\sigma}\right)(n-1)$ and $e^{r(\Delta+\bar{t}^k)} > e^{r\Delta}$, and hence $e^{r\Delta}/r < (n-1)/\sigma$, as desired.

□

In view of Lemma 3.1 and Lemma 3.4, we can concisely state results on the existence and uniqueness of SSPE and the associated surplus destruction in the following theorem:

Theorem 3.5. *In the game $G(\Delta, n, r, \sigma)$, there exists a unique SSPE. It is efficient if the inequality $e^{r\Delta}/r \geq (n-1)/\sigma$ holds, and it involves surplus destruction otherwise.*

Let us consider in more detail the relation between the rate of time preference and the two types of SSPE. To this end, define the twice continuously differentiable function $\mu(r) = e^{r\Delta}/r$ on the domain $(0, \infty)$ and consider its derivatives

$$\begin{aligned}\mu'(r) &= -r^{-2}e^{r\Delta} + r^{-1}\Delta e^{r\Delta},\\ \mu''(r) &= r^{-1}\Delta^2 e^{r\Delta} - r^{-2}\Delta e^{r\Delta} + 2r^{-3}e^{r\Delta} - r^{-2}\Delta e^{r\Delta}.\end{aligned}$$

We see that $\mu'(r) = 0$ if and only if $r = 1/\Delta$, and $\mu''(1/\Delta) = \Delta^3 e > 0$. Observe also that $\mu(r)$ grows without bound both in the limit as $r \to 0$ and in the limit as $r \to \infty$. It follows that the quasi–convex function $\mu(r)$ attains its unique minimum on the domain $(0, \infty)$ at the point $1/\Delta$, where it evaluates to $\mu(1/\Delta) = \Delta e$.

Invoking Theorem 3.5, we can now conclude that if Δ, n, and σ are such that $\Delta \geq (n-1)/(\sigma e)$, then the SSPE is efficient irrespective of the value of r. If, however, $\Delta < (n-1)/(\sigma e)$, then there is an intermediate range of values for r such that the SSPE involves surplus destruction. Recall that $\Delta > 0$ represents some minimal time lapse which occurs between bargaining rounds in the absence of any surplus destruction by the players, so that it makes sense to think of Δ as "small." In particular, we assume from now on that the model parameters Δ, n, and σ satisfy the inequality $\Delta < (n-1)/(\sigma e)$. As a result, there exists an open interval of values for r such that the SSPE involves surplus destruction for any r in this interval, and is efficient for any r outside of this interval. In particular, the said interval always contains $1/\Delta$. In the next section, we extend the foregoing analysis in order to quantify the amount of surplus destruction in an SSPE.

3.6 Surplus Destruction and Proposer Premium

The model in this paper abstracts away from any heterogeneity between the players. Hence, we focus on two sets of questions: First, we ask how much surplus (if any) the proposer destroys in equilibrium, and we investigate the relation between basic model parameters and the equilibrium level of surplus destruction. Second, we derive results on the extent of the premium which the proposer obtains in equilibrium.

More specifically, in Subsection 3.6.1 below, we show that the relation between the rate of time preference and the equilibrium level of surplus destruction is continuous, non–monotonic, and single–peaked. In Subsection 3.6.2, we derive an explicit upper bound on equilibrium surplus destruction, and we argue that this bound is "tight," that is, any level of surplus destruction strictly below this bound can occur in an SSPE. In Subsection 3.6.3, we provide a number of results pertaining to the proposer premium in the game at hand. Finally, in Subsection 6.4, we consider the extreme cases when the parameter σ is either very small or very large.

3.6.1 Continuity and Non–monotonicity of Surplus Destruction

For the purpose of this subsection, fix some Δ, n, and σ which are such that our earlier assumption $\Delta < (n-1)/(\sigma e)$ is satisfied. Given this parameter choice, let a map $\breve{t} : (0, \infty) \to [0, 1/\sigma]$ assign to each value of r the equilibrium level of t.

From Theorem 3.5, we know that $\breve{t}(r) = 0$ for values of r such that $e^{r\Delta}/r \geq (n-1)/\sigma$, and $\breve{t}(r) > 0$ for values of r such that $e^{r\Delta}/r < (n-1)/\sigma$. Together with the analysis of the quasi–convex function $\mu(r) = e^{r\Delta}/r$ in the previous section, this implies that there is an open interval $(\underline{r}, \overline{r})$ such that

$$\begin{aligned} \breve{t}(r) &= 0 \quad \text{if} \quad 0 < r \leq \underline{r}, \\ \breve{t}(r) &> 0 \quad \text{if} \quad \underline{r} < r < \overline{r}, \\ \breve{t}(r) &= 0 \quad \text{if} \quad \overline{r} \leq r. \end{aligned}$$

More in particular, for any r such that $\underline{r} < r < \overline{r}$, we know that $\breve{t}(r)$ is implicitly given by the solution to Eqn. (3.13), which we can equivalently rewrite as follows:

$$\frac{e^{r(\Delta+t)}}{r} = \left(\frac{(1-\sigma t)^2}{\sigma}\right)(n-1). \tag{3.16}$$

Observe that both sides of this equation are continuous in both r and t. This implies that $\breve{t}(r)$ is a continuous function on the interval $(\underline{r}, \overline{r})$. It is trivial that $\breve{t}(r)$ is also a continuous function on the intervals $(0, \underline{r}]$ and $[\overline{r}, \infty)$, where it is in fact constant. In order to show that $\breve{t}(r)$ is a continuous function on its entire domain, Note that Eqn. (3.16) reduces to $e^{r\Delta}/r = (n-1)/\sigma$ for $t = 0$, and that this equality implicitly defines the points $\underline{r}$ and $\overline{r}$.

Indeed, we see that a continuous function $\breve{t}(r)$ relates the rate of time preference to the time lapse chosen in an SSPE.

Analogously to the definition of $\mu(r)$ in the previous section, let us define for any $t \in (0, 1/\sigma]$ the function

$$\mu_t(r) = e^{r(\Delta+t)}/r.$$

Again analogously to the function $\mu(r)$ in the previous section, we see that $\mu_t(r)$ is a quasi–convex function, tends to infinity in the limit as $r \to 0$ and as $r \to \infty$, and attains the minimum at $r = 1/(\Delta + t)$, where it evaluates to $\mu_t(1/(\Delta + t)) = (\Delta + t)e$. The right–hand side of Eqn. (3.16) is constant in r. Hence, for any given values of Δ, n, and σ, we can distinguish the following three cases:

Case 1: If t is such that $(\Delta + t)e > \left(\frac{(1-\sigma t)^2}{\sigma}\right)(n-1)$, then there does not exist $r > 0$ which solves Eqn. (3.16). Consequently, the level of surplus destruction σt does not occur in the SSPE of any game $G(\Delta, n, r, \sigma)$, irrespective of r.

Case 2: If t is such that $(\Delta + t)e = \left(\frac{(1-\sigma t)^2}{\sigma}\right)(n-1)$, then $r = 1/(\Delta + t)$ is the unique solution to Eqn. (3.16). The SSPE of the game $G\left(\Delta, n, \frac{1}{\Delta+t}, \sigma\right)$ involves the level of surplus destruction σt.

Case 3: If t is such that $(\Delta + t)e < \left(\frac{(1-\sigma t)^2}{\sigma}\right)(n-1)$, then there exist two values r', r'' such that $0 < r' < \frac{1}{\Delta+t} < r''$ which solve Eqn. (3.16). Thus, the SSPE of the games $G(\Delta, n, r', \sigma)$ and $G(\Delta, n, r'', \sigma)$ involve the level of surplus destruction σt.

Eqn. (3.16) implies an upper bound on the level of surplus destruction which can occur in an SSPE. In the next subsection, we derive this bound explicitly. To conclude, we point out that the relation $\breve{t}(r)$ between the rate of time preference and the equilibrium level of surplus destruction is single–peaked. In order to see this, take $t^* > 0$ such that $e(\Delta + t^*) < \frac{(1-\sigma t^*)^2(n-1)}{\sigma}$. Then, there are r' and r'' such that $r' < \frac{1}{\Delta+t^*} < r''$ and the SSPE of $G(\Delta, n, r', \sigma)$ and of $G(\Delta, n, r'', \sigma)$ involve surplus destruction σt^*. Likewise, for sufficiently small $\varepsilon > 0$, there are r'_ε and r''_ε such that $r'_\varepsilon < \frac{1}{\Delta+t^*} < r''_\varepsilon$ and the SSPE of $G(\Delta, n, r'_\varepsilon, \sigma)$ and of $G(\Delta, n, r''_\varepsilon, \sigma)$ involve surplus destruction $\sigma(t^* - \varepsilon)$. Observe that $\mu_{t^*-\varepsilon}(r) < \mu_{t^*}(r)$ for all $r \in (0, \infty)$, and moreover, $\frac{(1-\sigma t^*+\sigma\varepsilon)^2(n-1)}{\sigma} > \frac{(1-\sigma t^*)^2(n-1)}{\sigma}$. Together with the continuity of both sides of Eqn. (3.16) in both r and t, this implies that $r'_\varepsilon < r'$ and $r''_\varepsilon > r''$.

3.6.2 An Explicit Bound on Surplus Destruction

In the previous subsection, we have argued that an SSPE with the level σt of surplus destruction exists if and only if the inequality

$$(\Delta + t)e \leq (n-1)(1-\sigma t)^2/\sigma$$

is satisfied. It can be rewritten in the following form:

$$(\sigma t)^2 - \left(2 + \frac{e}{n-1}\right)(\sigma t) + 1 - \Delta\left(\frac{\sigma e}{n-1}\right) \geq 0. \tag{3.17}$$

Hence, an SSPE with surplus destruction σt exists if and only if[8] the following inequality holds:

$$\sigma t \leq 1 + \frac{e}{2(n-1)} - \sqrt{\frac{1}{4}\left(2 + \left(\frac{e}{n-1}\right)\right)^2 - 1 + \Delta\left(\frac{\sigma e}{n-1}\right)}. \tag{3.18}$$

In the limit as $\Delta \downarrow 0$, the right–hand side converges to

$$\eta(n) = 1 + \frac{e}{2(n-1)} - \sqrt{\frac{1}{4}\left(2 + \left(\frac{e}{n-1}\right)\right)^2 - 1}.$$

Since $\Delta\left(\frac{\sigma e}{n-1}\right) > 0$, we have the following theorem:

Theorem 3.6. *In an SSPE with surplus destruction, we have $\sigma t < \eta(n)$.*

We want to show that the bound $\eta(n)$ is tight in the following sense: Any amount of surplus destruction strictly below $\eta(n)$ can occur in an SSPE if Δ is small and r is appropriately chosen. In order to show this, fix any $\sigma > 0$ and any $n = 2, 3, \ldots,$ and define

$$\widehat{\Delta}(\varepsilon) = \varepsilon\left(\frac{n-1}{e\sigma}\right)$$

for $\varepsilon > 0$. Plugging $\widehat{\Delta}(\varepsilon)$ into Ineq. (3.18), and considering Case 3 from the previous subsection, we obtain the following result:

Theorem 3.7. *Fix any $n \in \mathbb{Z}_+$, any $\sigma > 0$, and any $t < \eta(n)/\sigma$. There are $\varepsilon > 0$ sufficiently small and $r', r'' > 0$ such that the SSPE of the games $G(\widehat{\Delta}(\varepsilon), n, r', \sigma)$ and $G(\widehat{\Delta}(\varepsilon), n, r'', \sigma)$ involve surplus destruction σt.*

We stress that the level of surplus destruction can come arbitrarily close to the bound $\eta(n)$ irrespective of the value of the parameter σ, which indicates how much surplus is destroyed in order to prolong the time lapse between rounds by one unit. Intuitively, one might expect that a higher value of σ discourages surplus destruction. Our analysis reveals that this is only partially true: While a higher value of σ does narrow down the range of values of the rate of time preference for which surplus destruction occurs, it has no effect on the supremum of possible levels of surplus destruction in equilibrium.

Observe that $\eta(n+1) > \eta(n)$ for any $n \geq 2$: More surplus destruction becomes possible as the number of players increases. The table shows the upper bound $\eta(n)$ on equilibrium surplus destruction for various values of n.

[8] Note that Eqn. (3.17) is also satisfied if $\sigma t \geq 1 + \frac{e}{2(n-1)} + \sqrt{\frac{1}{4}\left(2 + \left(\frac{e}{n-1}\right)\right)^2 - 1 + \Delta\left(\frac{\sigma e}{n-1}\right)}$, but this case is irrelevant here because $\sigma t \leq 1$ by assumption.

n	2	3	4	5	6	7	...	1000
$\eta(n)$	0.22	0.33	0.40	0.45	0.49	0.52	...	0.95

Since we are considering unanimity bargaining, it seems natural to have in mind only a "small" number of players. Nevertheless, it might be noteworthy that $\eta(n)$ cannot be bounded away from one in the limit as the number of players becomes arbitrarily large.

3.6.3 Results on the Proposer Premium

This subsection deals with the proposer's payoff and the extent of her first–mover advantage. First, we compare the proposer's payoff in the game at hand to that in a bargaining game with fixed timing. We show that the endogenous choice of the time lapse through surplus destruction benefits the proposer. Second, we compare the extent of the proposer premium for different values of the rate of time preference. We show that, despite the non–monotonic influence of the rate of time preference on the level of surplus destruction, it is still true that the proposer premium is strictly monotonically increasing in the rate of time preference.

In the previous section, we have introduced the notion of the "fixed timing allocation." This is the equilibrium allocation which would result in a unanimity bargaining game in which players cannot destroy surplus and thus cannot affect the time lapse $\Delta > 0$ which occurs between bargaining rounds. Intuitively, the bargaining game which we consider gives the proposer an additional source of power by allowing her to destroy surplus and thereby prolong the time lapse between rounds. This has two effects: On the one hand, the total amount of surplus allocated to the players is smaller, but on the other hand, the proposer's share of it is greater. It is not a priori clear whether the proposer is better off in the bargaining game under consideration here or in the bargaining game with fixed timing. We address this question in the following lemma.

Lemma 3.5. *In an SSPE with surplus destruction, the proposer receives more than her fixed timing allocation.*

The proof of Lemma 3.5 is provided in Appendix C.

Combined with the immediate agreement property, this result says that introducing the possibility of surplus destruction into a unanimity bargaining game is beneficial for the initial proposer only, but harmful for all other players. More specifically, all players other than the proposer are harmed in two different ways: The surplus effectively allocated is smaller than in the fixed timing allocation, and the responding players' share of it is also smaller.

Note that the above lemma is not in contradiction to our earlier finding that there is a tight bound $\eta(n)$ on surplus destruction which approaches one arbitrarily closely in the limit as n grows without bound. The reason is that the proposer's fixed timing allocation is given by $\frac{1}{1+(n-1)e^{-r\Delta}}$, which clearly goes to zero for sufficiently large n.

Even the fixed timing allocation always involves an advantage for the proposer. With homogeneous players, this readily implies that the proposer obtains more than an equal share of the surplus. Consequently, this must also be true in an SSPE of the game at hand, whence the following corollary.

Corollary 3.1. *In an SSPE, the proposer receives a greater share than* $1/n$ *of the surplus.*

The idea behind the model under consideration is that a proposer can consolidate her advantageous position by destroying some surplus. As a result, the extent of the benefit from being the proposer becomes endogenous. We investigate how much a player benefits from being the proposer in equilibrium. In what follows, we will mean by the *proposer premium* the ratio of the proposer's payoff to the payoff of a responding player. Thus, the proposer premium in the fixed timing allocation (and thus in an efficient SSPE) is $e^{r\Delta}$, while the proposer premium in an SSPE with surplus destruction is given by Eqns. (3.11)–(3.12).

Under the fixed timing allocation, there is a monotonic relation between the rate of time preference and the proposer premium. We are now going to show that this monotonicity is preserved in SSPE of the game at hand.

Theorem 3.8. *For any given* $\Delta, n,$ *and* σ *with* $\Delta < (n-1)/(\sigma e)$, *the proposer premium is strictly monotonically increasing in the rate of time preference.*

The proof of Theorem 3.8 is relegated to Appendix C.

In the literature on Rubinstein bargaining, it is a standard result that the proposer premium vanishes in the limit as players become arbitrarily patient. We argue that this is also true in the bargaining game under consideration here. In order to see this, fix some values of Δ, n, and σ, and consider a sequence of bargaining games with these parameters along which the rate of time preference $r > 0$ goes to zero. For sufficiently small r, we know from Theorem 3.5 that SSPE is efficient. Moreover, an efficient SSPE always leads to the fixed timing allocation. The ratio of the proposer's payoff to a responding player's payoff is $e^{r\Delta}$ in the fixed timing allocation. In the limit as $r \to 0$, this term converges to one, and so the allocation converges to the equal split $(1/n, \ldots, 1/n)$.

Corollary 3.2. *Fix* $\Delta, n,$ *and* σ, *and consider a sequence of bargaining games* $\{G(\Delta, n, r_\kappa, \sigma)\}_{\kappa=0,1,\ldots}$ *such that* $0 < r_{\kappa+1} < r_\kappa$ *for all* κ, *and* $\lim_{\kappa\to\infty} r_\kappa = 0$. *Let* ρ_κ *be the proposer premium in the SSPE of the game* $G(\Delta, n, r_\kappa, \sigma)$. *Then,* $\lim_{\kappa\to\infty} \rho_\kappa = 1$.

Throughout the present paper, we use the bargaining game with fixed timing as a benchmark. This game belongs to the class of unanimity bargaining games in the tradition of Rubinstein. This class of games has been extensively studied in the literature. Many important results in this literature pertain to the limit case where the exogenously given discount factor goes to one. That is, while the "bargaining friction" is the fundamental source of power in Rubinstein bargaining, the case which attracts special interest in the literature is the case where exactly this friction vanishes. A vanishing discount factor in Rubinstein bargaining can be interpreted in two ways: Discounting may vanish because players become arbitrarily patient, or because the time lapse between proposals become arbitrarily small so that the bargaining process features "frequent offers."

Our paper differs from the existing literature in that respect. Rather than considering the case where discounting becomes negligible, we make the extent of discounting endogenous. Moreover, arbitrarily patient players and arbitrarily frequent offers are not equivalent in our model. In particular, the above corollary says that, as in Rubinstein bargaining, discounting becomes negligible as players become sufficiently patient. However, we have shown in Theorem 3.7 above that discounting does not vanish as the exogenously given minimal time lapse Δ vanishes. On the contrary, players can be enticed to "compensate" a smaller value of Δ by destroying more surplus.

3.6.4 Varying the Proposer's Gain from Surplus Destruction

In the present section, we have mainly considered the effect of the rate of time preference on the proposer premium and on the level of surplus destruction. Thus, we have obtained results that can be contrasted with previous findings in the non–cooperative bargaining literature.

Alternatively, one might also want to consider how the SSPE in our model responds to changes in the parameter σ. To be more precise, one might want to hold the model parameters Δ, n, and r fixed, and conduct a comparative statics analysis with regard to σ.

Recall the assumption that the share σt of the surplus is destroyed if the proposer chooses to prolong the time lapse by t. Thus, the parameter σ measures how detrimental rent–seeking is for the surplus. Equivalently, we can also say that a high σ indicates that the proposer's gain from surplus destruction is low, and vice versa.

Two cases deserve special attention: First, suppose that σ is very large. Verbally, this means that the proposer is able to destroy surplus, but there is (almost) no effect on her bargaining power relative to the responding players. In this case, it is intuitive that it makes no sense for the proposer to destroy surplus. Indeed, it follows from Theorem 3.5 that SSPE is efficient if σ is sufficiently large.

Second, suppose that σ is close to zero. Verbally, this means that the proposer can make powerful threats against the responding players while losing (almost) no surplus. In that case, it is intuitive that the proposer wants to choose a very long time lapse. Indeed, it is easy to see that the equilibrium choice of the time lapse is a decreasing function

$\widehat{t}(\sigma)$. However, it seems ambiguous how the equilibrium level of surplus destruction $\sigma\widehat{t}(\sigma)$ changes in the limit as σ goes to zero. We claim that equilibrium surplus destruction vanishes as σ becomes sufficiently small.[9] We conclude that, for σ sufficiently small, (almost) no surplus is destroyed, and (almost) the entire surplus goes to the proposer. Earlier in this section, we have found a non–monotonic relationship between r and the equilibrium level of surplus destruction. In this subsection, we have also observed a non–monotonic relationship between σ and the equilibrium level of surplus destruction.

3.7 Conclusion

We have revisited a unanimity bargaining game in the tradition of Rubinstein (1982). In this kind of non–cooperative bargaining game, the proposer derives power from the prospect of a costly delay which would occur if her proposal was rejected. We have augmented this canonical model of non–cooperative bargaining by making the length of the time lapse between proposals endogenous. The proposer can choose to sacrifice some surplus in order to prolong this time lapse. Surplus destruction allows the proposer to become more deeply entrenched in her privileged position.

We have demonstrated the existence of a unique SSPE and obtained sharp predictions for the level of surplus destruction as well as for the allocation. The model offers insight on a number of issues:

First, one crucial question is whether a proposer does or does not make use of the possibility of surplus destruction in equilibrium. For different parameter scenarios, SSPE may be efficient, but may also involve the destruction of a substantial share of the surplus. If the SSPE is efficient, then our model replicates results familiar from the study of the canonical unanimity bargaining game with fixed timing. In particular, it is true that the proposer receives nearly the entire surplus when players are sufficiently impatient. It is also true that the surplus is split nearly evenly among all players when they are sufficiently patient. For intermediate degrees of patience, surplus destruction occurs as an equilibrium phenomenon, and so our predictions depart from the canonical ones. In particular, we establish a tight upper bound on the level of surplus destruction. This upper bound depends only on the number of players. Given the number of players, any level of surplus destruction strictly below the aforementioned bound can be attained for some choice of the rate of time preference. Therefore, while the number of players determines the tight upper bound on surplus destruction, the actual amount of surplus destruction is determined by the players' patience. In the present paper, we have assumed homogeneous

[9] In order to see this, suppose by way of contradiction that there is a lower bound $b > 0$ on equilibrium surplus destruction in the limit as σ tends to zero. Thus, $\widehat{t}(\sigma) \geq b/\sigma$. Moreover, $e^{r(\Delta+\widehat{t}(\sigma))} \leq \left(\frac{r}{\sigma}\right)(1-b)^2(n-1)$. Combining these two inequalities yields $\sigma e^{r(\Delta+b/\sigma)} \leq r(n-1)(1-b)^2$. Observe that the right–hand side is independent of σ, while the left–hand side tends to infinity as σ goes to zero from above. This is the desired contradiction.

players in order to focus on an analysis of surplus destruction and the proposer premium. In the literature on bargaining games with fixed timing, it is known, however, that players' relative patience is an important source of bargaining power. A more patient player can secure a greater share of the surplus. One pertinent extension of our model would be to study players which are heterogeneous with regard to the rate of time preference.

Second, we show that the bargaining process under consideration tends to become less efficient when the number of players increases. The driving force behind this result is the bargaining protocol, which allows the rejecting player to make the next proposal. Consequently, the proposer uses surplus destruction to simultaneously discourage rejections by all other players.

Third, one might intuitively expect that surplus destruction is discouraged when σ is high, and thus the proposer's gain from surplus destruction is low. It turns out that this is only partially true: On the one hand, a lower gain from surplus destruction does narrow down the range of values for the rate of time preference on which surplus destruction occurs in equilibrium. On the other hand, it is somewhat surprising that the supremum of surplus destructions which can be obtained in equilibrium is entirely unaffected by the proposer's gain from surplus destruction.

Fourth, one might wonder whether the possibility of rent–seeking through surplus destruction harms all players compared to canonical Rubinstein bargaining. In particular, one might have expected that the equilibrium allocation obtained in the bargaining game in this paper is Pareto–dominated by the equilibrium allocation in a bargaining game with fixed timing. It turns out that this is not the case. In particular, if the SSPE in our game is efficient, then it involves the same surplus allocation as the equilibrium of a bargaining game with fixed timing. Moreover, if the SSPE of our game involves surplus destruction, then only the responding players are worse off than in a bargaining game with fixed timing; the proposer is better off than she would be in the equilibrium of a bargaining game with fixed timing.

Finally, we show that the proposer benefits unambiguously from a higher rate of time preference. Our model inherits this feature from standard unanimity bargaining games with fixed timing.

Appendix A

The purpose of this appendix is to show that (3.1)–(3.3) are necessary conditions for an SSPE. Moreover, we will show that any SSPE leads to immediate agreement in any subgame. In order to establish these claims, several auxiliary results are needed. The basic line of reasoning in this appendix is similar to that used in previous work by Banks and Duggan (2000) and Britz et al. (2010, 2014). Compared to those papers, however, the endogenous determination of the time lapse adds a number of complications to the argument.

The first step is to derive some conditions on the players' acceptance sets.

Lemma 3.6. *Suppose that $(\theta^k, t^k, A^k(t))_{k \in N}$ is an SSPE. Let $i \in N$ and $\widehat{t} \in [0, 1/\sigma]$.*

1. *If a proposal $v \in \mathbb{R}^n_+$ satisfies the inequality*

$$v_i < (1 - \sigma\widehat{t}\,)e^{-r(\Delta+\widehat{t}\,)}\gamma_i((\theta^k, t^k, A^k(t))_{k \in N}),$$

   then $v \notin A(\,\widehat{t}\,)$.

2. *It holds that*

$$A^i(\,\widehat{t}\,) \supset \{v \in \mathbb{R}^n_+ | v_j > (1 - \sigma\widehat{t}\,)e^{-r(\Delta+\widehat{t}\,)}\gamma_j((\theta^k, t^k, A^k(t))_{k \in N}),\ \forall\, j = i, \dots, n\}.$$

Proof. *Part 1.* Consider a history where the current proposer has chosen $\widehat{t}$ and made the proposal v. Suppose that $v \in A(\,\widehat{t}\,)$: All players accept v. Consider a unilateral deviation by Player i which consists of rejecting v. This gives Player i a payoff of $(1 - \sigma\widehat{t}\,)e^{-r(\Delta+\widehat{t}\,)}\gamma_i((\theta^k, t^k, A^k(t))_{k \in N})$, establishing the first part of the lemma.

Part 2. The proof is by backward induction. First consider a history at which the current proposer has chosen $\widehat{t}$ and made a proposal v, the Players $1, \dots, n-1$ have all accepted the proposal v, and now it is Player n's turn to decide whether to accept or reject v. Note that v is going to be implemented if and only if Player n accepts it. Thus, Player n receives a payoff v_n if he accepts, or a payoff of $(1 - \sigma\widehat{t}\,)e^{-r(\Delta+\widehat{t}\,)}\gamma_n((\theta^k, t^k, A^k(t))_{k \in N})$ if he rejects. Indeed, it is optimal for Player n to accept if $v_n > (1 - \sigma\widehat{t}\,)e^{-r(\Delta+\widehat{t}\,)}\gamma_n((\theta^k, t^k, A^k(t))_{k \in N})$. Now consider the history where the current proposer has chosen $\widehat{t}$ and made the proposal v, then the Players $1, \dots, n-2$ have all accepted v, and it is the turn of Player $n-1$ to accept or reject v. If $v_n > (1 - \sigma\widehat{t}\,)e^{-r(\Delta+\widehat{t}\,)}\gamma_n((\theta^k, t^k, A^k(t))_{k \in N})$, then we have already shown that Player n is going to accept the proposal v, and hence v is going to be implemented if and only if Player $n-1$ accepts it. Repeating the previous argument inductively for Players $n-1, n-2, \dots, 1$ establishes the second part of the lemma. □

One important step is to show that, in an SSPE, every subgame will ultimately end in an agreement. In particular, this means that no proposer ever decides to end the game by destroying the entire surplus, and no perpetual disagreement can occur on the equilibrium path of play. As a first step towards this result, Lemma 3.7 below claims that, in an SSPE, there is at least one player whose proposal is unanimously accepted.

Lemma 3.7. *If $(\theta^k, t^k, A^k(t))_{k\in N}$ is an SSPE, then there is $i \in N$ such that $\theta^i \in A(t^i)$, and $t^i < 1/\sigma$.*

Proof. Suppose by way of contradiction that $(\theta^k, t^k, A^k(t))_{k\in N}$ is an SSPE, and that for every $k \in N$, it holds either that $\theta^k \notin A(t^k)$, or that $t^k = 1/\sigma$. Then, on the path of equilibrium play, either there is perpetual disagreement, or the entire surplus is destroyed. In either case, all players receive zero payoffs. Hence, in the supposed SSPE, we have $\gamma_k((\theta^k, t^k, A^k(t))_{k\in N}) = 0$ for all $k \in N$. Due to (the second part of) Lemma 3.6 above, this implies that $\mathbb{R}^n_{++} \subset A(0)$. Take a Player i, and consider a history at which Player i chooses a proposal and a time lapse. Suppose that he deviated from $(\theta^k, t^k, A^k(t))_{k\in N}$ by choosing $(\tilde{\theta}^i, \tilde{t}^i)$ with $\tilde{\theta}^i = (1/n, \dots, 1/n)$ and $\tilde{t}^i = 0$. Since $\mathbb{R}^n_{++} \subset A(0)$, the proposal $\tilde{\theta}^i$ is unanimously accepted, and so the deviation is profitable for Player i.

□

We have shown that at least one player makes an acceptable proposal in an SSPE. Now we consider a proposal which is unanimously accepted, and show that such a proposal makes each player other than the proposer indifferent between acceptance and rejection. Moreover, such a proposal distributes the entire share of the surplus which is not destroyed.[10] It is a standard result in unanimity bargaining that the proposer "extracts all surplus" from the other players by making them exactly indifferent between acceptance and rejection. Lemma 3.8 below replicates this finding for the game under consideration here.

Lemma 3.8. *Suppose that $(\theta^k, t^k, A^k(t))_{k\in N}$ is an SSPE, and $\theta^i \in A(t^i)$, as well as $t^i < 1/\sigma$. Then, we have*

$$\begin{aligned}
\theta^i_j &= (1-\sigma t^i)e^{-r(\Delta+t^i)}\gamma_j((\theta^k, t^k, A^k(t))_{k\in N}),\ \forall i \in N,\ \forall j \in N\setminus\{i\},\\
\theta^i_i &= 1-\sigma t^i - \sum_{j\in N\setminus\{i\}} \theta^i_j,\ \forall i \in N.
\end{aligned}$$

[10] In the existing literature on Rubinstein bargaining, it is common to assume that the proposer is allowed to make a proposal which leaves some surplus unallocated. It is then easily shown that such a proposal is never made in an SSPE. In the model at hand, the proposer destroys some share of the surplus, and allocates some share to each player. The rules of the game do not require that the shares allocated to players and the share destroyed sum up to one. However, we show that they do sum up to one in an SSPE.

Proof. *Part 1.* Due to (the first part of) Lemma 3.6 above, $\theta^i \in A(t^i)$ implies

$$\theta^i_j \geq (1 - \sigma t^i) e^{-r(\Delta + t^i)} \gamma_j((\theta^k, t^k, A^k(t))_{k \in N})$$

for all $i, j \in N$. Suppose by way of contradiction that $(\theta^k, t^k, A^k(t))_{k \in N}$ is an SSPE, but there are Players $i \in N$ and $j \in N \setminus \{i\}$ such that

$$\theta^i_j > (1 - \sigma t^i) e^{-r(\Delta + t^i)} \gamma_j((\theta^i, t^i, A^i(t))_{i \in N}).$$

Let $\tilde{\theta}^i_j = \theta^i_j - \varepsilon$ and $\tilde{\theta}^i_i = \theta^i_i + \varepsilon$ while $\tilde{\theta}^i_k = \theta^i_k$ for all $k \in N \setminus \{i, j\}$. For $\varepsilon > 0$ sufficiently small, it holds that $\tilde{\theta}^i \in A(t^i)$. Hence, making the proposal $\tilde{\theta}^i$ instead of θ^i is a profitable deviation for Player i.

Part 2. Since equilibrium proposals are required to be feasible, the inequality $\theta^i_i \leq 1 - \sigma t^i - \sum_{j \in N \setminus \{i\}} \theta^i_j$ holds in an SSPE. We have to show that it must hold with equality. Suppose not. Then, a proposal $\tilde{\theta}^i$ defined by $\tilde{\theta}^i_i = \theta^i_i + \varepsilon$ and $\tilde{\theta}^i_j = \theta^i_j$ for $j \in N \setminus \{i\}$ would be acceptable and still satisfy the inequality $1 - \sigma t^i \geq \sum_{k \in N} \tilde{\theta}^i_k$ for $\varepsilon > 0$ sufficiently small.
$\square$

The next step is to show that a proposer obtains a strictly positive payoff in an SSPE.

Lemma 3.9. *If $(\theta^k, t^k, A^k(t))_{k \in N}$ is an SSPE, then it holds that $\gamma_j((\theta^k, t^k, A^k(t))_{k \in N}) > 0$ for all $j \in N$.*

Proof. Suppose by way of contradiction that $(\theta^k, t^k, A^k(t))_{k \in N}$ is an SSPE, but there is $j \in N$ such that $\gamma_j((\theta^k, t^k, A^k(t))_{k \in N}) = 0$.

Step 1. Suppose that there is some $t' \in [0, 1/\sigma]$ such that

$$(1 - \sigma t') e^{-r(\Delta + t')} \sum_{k \in N} \gamma_k((\theta^k, t^k, A^k(t))_{k \in N}) < 1 - \sigma t'.$$

Then, let θ' be a proposal given by $\theta'_k = (1 - \sigma t') e^{-r(\Delta + t')} \gamma_k((\theta^k, t^k, A^k(t))_{k \in N}) + \varepsilon$ for every $k \in N \setminus \{j\}$, and $\theta'_j = \varepsilon > 0$. Due to the above inequality, the proposal θ' is feasible for $\varepsilon > 0$ sufficiently small. Moreover, due to (the second part of) Lemma 3.6 above, it holds that $\theta' \in A(t')$. So, Player j can profitably deviate from the supposed SSPE by choosing (θ', t'), thus receiving the share $\theta'_j = \varepsilon > 0$ instead of the share $\gamma_j((\theta^k, t^k, A^k(t))_{k \in N}) = 0$. We have now shown by contradiction that

$$(1 - \sigma t) e^{-r(\Delta + t)} \sum_{k \in N} \gamma_k((\theta^k, t^k, A^k(t))_{k \in N}) \geq 1 - \sigma t, \ \forall t \in [0, 1/\sigma].$$

Step 2. We have seen in Lemma 3.7 that there is a Player $i \in N$ such that $t^i < 1/\sigma$ and $\theta^i \in A(t^i)$. Due to (the first part of) Lemma 3.6 above, this implies the inequality

$$(1-\sigma t^i)e^{-r(\Delta+t^i)} \sum_{k\in N} \gamma_k((\theta^k, t^k, A^k(t))_{k\in N}) \leq 1-\sigma t^i.$$

But in Step 1 above, we have shown that the reverse inequality holds for all $t \in [0, 1/\sigma]$, and in particular for t^i. Therefore,

$$(1-\sigma t^i)e^{-r(\Delta+t^i)} \sum_{k\in N} \gamma_k((\theta^k, t^k, A^k(t))_{k\in N}) = 1-\sigma t^i.$$

Step 3. Due to Lemma 3.8, we have that $\theta^i_k = (1-\sigma t^i)e^{-r(\Delta+t^i)}\gamma_k((\theta^k, t^k, A^k(t))_{k\in N})$ for every $k \in N \setminus \{i\}$. Moreover, since the proposal θ^i is accepted, it holds by definition of $\gamma_i(.)$ that $\theta^i_i = \gamma_i((\theta^k, t^k, A^k(t))_{k\in N})$. Now the equality derived in Step 2 above can be written as

$$(1-\sigma t^i)e^{-r(\Delta+t^i)}\theta^i_i + \sum_{k\in N\setminus\{i\}} \theta^i_k = 1-\sigma t^i.$$

But due to Lemma 3.8, it holds that $\theta^i_i + \sum_{k\in N\setminus\{i\}} \theta^i_k = 1-\sigma t^i$. Hence, we conclude that $(1-\sigma t^i)e^{-r(\Delta+t^i)} = 1$, or, equivalently, $(1-\sigma t^i) = e^{r(\Delta+t^i)}$. But since $t^i \geq 0$ and $\Delta > 0$ as well as $r > 0$, we have the chain of inequalities $1-\sigma t^i \leq 1 < e^{r\Delta} \leq e^{r(\Delta+t^i)}$, a contradiction.

□

The rules of the bargaining game under consideration allow for perpetual disagreement as well as for the destruction of the entire surplus. One important implication of Lemma 3.9 above is that neither of these two outcomes can occur in an SSPE. To see this, suppose that either perpetual disagreement or the destruction of the entire surplus occurs in some subgame $G^i(\Delta, n, r, \sigma)$. This results in zero payoffs for all players, and in particular for the initial proposer in the subgame, thus contradicting Lemma 3.9. We have thus obtained the following corollary.

Corollary 3.3. *In an SSPE $(\theta^k, t^k, A^k(t))_{k\in N}$, a unanimous agreement is reached in finite time in every subgame $G^i((\theta^k, t^k, A^k(t))_{k\in N})$, for every $i \in N$.*

Since no player destroys the entire surplus, it follows that the "implicit discount factor" given by $(1-\sigma t^i)e^{-r(\Delta+t^i)}$ is strictly positive for every $i \in N$. Hence, Lemma 3.8 and Lemma 3.9 readily imply that any proposal θ^i which is accepted in an SSPE is strictly positive in all components. This property is crucial for the proof of Lemma 3.10 below, which claims that, in an SSPE, agreement is reached immediately in every subgame.

Lemma 3.10. *In an SSPE $(\theta^k, t^k, A^k(t))_{k\in N}$, it holds that $\theta^i \in A(t^i)$ for every $i \in N$, so that agreement is reached immediately in every subgame $G^i((\theta^k, t^k, A^k(t))_{k\in N})$.*

Proof. Let $(\theta^k, t^k, A^k(t))_{k\in N}$ be an SSPE. Now suppose that there is a Player $j \in N$ such that $\theta^j \notin A(t^j)$. We have shown before that there is no perpetual disagreement in an SSPE, and no player destroys the entire surplus. Thus, on the equilibrium path of play, the rejection of θ^j is followed by an agreement on the proposal θ^i of some Player i after a delay of $\hat{\tau} > 0$. Consider a deviation by Player j, who chooses $(\tilde{\theta}^j, \tilde{t}^j) = (\theta^i, t^i)$ instead of (θ^j, t^j). Since $\theta^i \in A(t^i)$, this deviation leads to immediate acceptance of the proposal $\tilde{\theta}^j$, and Player j receives the share $\tilde{\theta}^j_j = \theta^i_j$. But in the supposed SSPE, Player j receives that same share θ^i_j only after a delay of $\hat{\tau}$. Moreover, combining Lemma 3.9 with Lemma 3.8, we see that $\theta^i_j > 0$. Indeed, avoiding the delay $\hat{\tau}$ is strictly profitable, and so we have constructed a profitable deviation for Player j, which completes the proof.

□

Since every player's proposal is accepted in an SSPE, we have

$$\theta^i_i = \gamma_i((\theta^k, t^k, A^k(t))_{k\in N}) > 0$$

for every $i \in N$. Now Lemma 3.8 implies that in every SSPE,

$$\theta^i_j = (1-\sigma t^i)e^{-r(\Delta+t^i)}\theta^j_j > 0,$$

for every $i \in N$ and $j \in N \setminus \{i\}$. We have argued before that every proposal which is accepted in an SSPE is strictly positive in all components. Combined with Lemma 3.10, this yields the following corollary.

Corollary 3.4. *In an SSPE, every player makes a proposal which is strictly positive in all components.*

In Lemma 3.6, we have derived some conditions that the acceptance sets have to satisfy in an SSPE. We are now ready to strengthen that result. More specifically, Lemma 3.11 shows that only one specific acceptance set is consistent with the necessary conditions for an SSPE.

Lemma 3.11. *Suppose that $(\theta^k, t^k, A^k(t))_{k\in N}$ is an SSPE. Let $i \in N$ and $\widehat{t} \in [0, 1/\sigma]$. Then,*

$$A^i(\widehat{t}) = \{v \in \mathbb{R}^n_+ | v_j \geq (1-\sigma\widehat{t})e^{-r(\Delta+\widehat{t})}\gamma_j((\theta^k, t^k, A^k(t))_{k\in N}),\ \forall j = i, \ldots, n\}.$$

Proof. From Lemma 3.6, we know that

$$A^i(\widehat{t}) \supset \{v \in \mathbb{R}^n_+ | v_j > (1-\sigma\widehat{t})e^{-r(\Delta+\widehat{t})}\gamma_j((\theta^k, t^k, A^k(t))_{k\in N}),\ \forall j = i, \ldots, n\}.$$

According to Lemma 3.8, any SSPE proposal gives responding players exactly the amount that makes them indifferent with rejecting the proposal. Moreover, according to Lemma

3.10, an SSPE proposal is unanimously accepted. Hence, it follows that

$$A^i(\widehat{t}\,) \supset \{v \in \mathbb{R}^n_+ | v_j \geq (1-\sigma\widehat{t}\,)e^{-r(\Delta+\widehat{t}\,)}\gamma_j((\theta^k, t^k, A^k(t))_{k\in N}),\ \forall j = i, \ldots, n\}.$$

In order to complete the proof of the lemma, we have left to show that

$$A^i(\widehat{t}\,) \subset \{v \in \mathbb{R}^n_+ | v_j \geq (1-\sigma\widehat{t}\,)e^{-r(\Delta+\widehat{t}\,)}\gamma_j((\theta^k, t^k, A^k(t))_{k\in N}),\ \forall j = i, \ldots, n\}.$$

Suppose by way of contradiction that there is a time lapse $\widehat{t}$, a proposal v, and a Player i such that $v_{j'} < (1-\sigma\widehat{t}\,)e^{-r(\Delta+\widehat{t}\,)}\gamma_{j'}((\theta^k, t^k, A^k(t))_{k\in N})$ for some $j' \in \{i, \ldots, n\}$, but $v \in A^i(\widehat{t}\,)$. Due to (the first part of) Lemma 3.6, we have that $v \notin A(\widehat{t}\,)$. Consider the history where Player i decides whether to accept or reject v. This history is only reached if all players $1, \ldots, i-1$ have accepted v. Thus, the fact that $v \notin A(\widehat{t}\,)$ implies that there is a player $j'' \in \{i+1, \ldots, n\}$ such that $v \notin A^{j''}(\widehat{t}\,)$. If Player i accepts v, then v will subsequently be rejected by Player j'', and according to Lemma 3.8, this leads to a payoff of $(1-\sigma\widehat{t}\,)(1-\sigma t^{j''})e^{-r(\Delta+\widehat{t}\,)}e^{-r(\Delta+t^{j''})}\gamma_i(.)$ for Player i. But if Player i rejects v, then the resulting payoff is $(1-\sigma\widehat{t}\,)e^{-r(\Delta+\widehat{t}\,)}\gamma_i(.)$ Due to the assumption that $r > 0$ and $\Delta > 0$, the latter is strictly greater than the former, and the proof is complete.

□

Lemma 3.12. *Suppose that $(\theta^k, t^k, A^k(t))_{k\in N}$ is a profile of stationary strategies which satisfies Eqns. (3.1)–(3.4). Then, for any Player $i \in N$, and for any $\widehat{t} \in [0, 1/\sigma]$, the acceptance set $A^i(\widehat{t}\,)$ as specified in Eqn. (3.3) is optimal.*

Proof. Consider a proposal $v \in A^i(\widehat{t}\,)$. If Player i rejects v, then he obtains $(1-\sigma\widehat{t})e^{-r(\Delta+\widehat{t}\,)}\theta^i_i$, which is no greater than v_i for any $v \in A^i(\widehat{t}\,)$. Clearly, this deviation is not profitable. Now consider a proposal $v \notin A^i(\widehat{t}\,)$, and suppose that Player i accepts v. There are two cases: Either v is unanimously accepted, or it is subsequently rejected by some Player $j \in N \setminus \{i\}$. Suppose first that v is unanimously accepted. Then, Player i receives $v_i < (1-\sigma\widehat{t})e^{-r(\Delta+\widehat{t}\,)}\theta^i_i$, so the deviation is not profitable. Suppose next that, following Player i's acceptance, the proposal v is rejected by some Player $j \in N \setminus \{i\}$. In that case, Player i receives the payoff $(1-\sigma\widehat{t})e^{-r(\Delta+\widehat{t}\,)}\theta^i_j$. If he had not deviated, he would have rejected v himself and earned the payoff $(1-\sigma\widehat{t})e^{-r(\Delta+\widehat{t}\,)}\theta^i_i$. Since r and Δ are strictly positive, it follows from Eqn. (3.2) that $\theta^i_i > \theta^i_j$, indeed, the deviation for Player i is not profitable, and the proof is complete. □

To conclude, we see from Lemma 3.8 that Eqns. (3.1)–(3.2) are necessary conditions for an SSPE. Moreover, Lemma 3.11 shows that Eqn. (3.3) is a necessary condition for an SSPE. Lemma 3.10 establishes the immediate agreement property in every subgame. Lemma 3.12 shows that the four conditions (3.1)–(3.4) imply that the accept/reject decisions are optimal.

Appendix B

Proof of Lemma 3.3. Suppose that $(\theta^k, t^k)_{k\in N}$ is an SSPE. We show first that $t^i > 0$ for at least one $i \in N$ implies $(t^1, \ldots, t^n) \gg 0$. To see this, suppose by way of contradiction that $(\theta^k, t^k)_{k\in N}$ is an SSPE, and $t^i > 0$ but $t^j = 0$ for some $j \in N \setminus \{i\}$. Using Eqns. (3.1)–(3.2), it follows that $\theta^j_j + e^{-r\Delta}\sum_{k\in N\setminus\{j\}}\theta^k_k = 1$, which can be rewritten as

$$(1 - e^{-r\Delta})\theta^j_j + e^{-r\Delta}\sum_{k\in N}\theta^k_k = 1.$$

We consider a unilateral deviation by Player i from the supposed SSPE. Under this deviation, Player i chooses a zero time lapse instead of $t^i > 0$, and instead of the proposal θ^i, she makes a proposal $\tilde{\theta}^i$ defined as follows:

$$\begin{aligned}
\tilde{\theta}^i_k &= e^{-r\Delta}\theta^k_k,\ \forall k \in N \setminus \{i\},\\
\tilde{\theta}^i_i &= e^{-r\Delta}\theta^i_i + (1 - e^{-r\Delta})\theta^j_j.
\end{aligned}$$

We see that $\sum_{k\in N}\tilde{\theta}^i_k = 1$, so that the pair $(\tilde{\theta}^i, 0)$ is feasible. Due to Eqns. (3.2)–(3.3), we have $\tilde{\theta}^i \in A(0)$, so the proposal $\tilde{\theta}^i$ is unanimously accepted. We see that the deviation is profitable for Player i if $e^{-r\Delta}\theta^i_i + (1-e^{-r\Delta})\theta^j_j > \theta^i_i$, or, equivalently, if $\theta^j_j > \theta^i_i$. Indeed, due to Theorem 3.2, we know that $e^{-r\Delta}\sum_{k\in N\setminus\{j\}}\theta^k_k \leq \frac{\sigma}{\sigma+r}$. By Eqns. (3.1)–(3.2), it follows that $\theta^j_j \geq \frac{r}{\sigma+r}$. But from Eqn. (3.14), we have that $\theta^i_i = \frac{r(1-\sigma t^i)^2}{\sigma+r(1-\sigma t^i)} < \frac{r}{\sigma+r}$, and thus $\theta^i_i < \theta^j_j$. Indeed, we have constructed a profitable deviation for Player i and thus obtained a contradiction.

Now we have left to show that if $(t^1, \ldots, t^n) \gg 0$, then $t^1 = \ldots = t^n$. To this end, let us define a function

$$\psi(t) = \frac{r(1-\sigma t)^2 + \sigma e^{r(\Delta+t)}}{\sigma + r(1-\sigma t)}.$$

Suppose that $(\theta^k, t^k)_{k\in N}$ is an SSPE and $(t^1, \ldots, t^n) \gg 0$. Thus, Eqns. (3.8) and (3.10) hold for all $i \in N$. Summing up these two equations we find that $\psi(t^i) = \sum_{k\in N}\theta^k_k$ for every $i \in N$. We have left to show that the function $\psi(t)$ is strictly monotonic on the relevant interval $t \in (0, 1/\sigma]$. Indeed, consider the first–order derivative

$$\begin{aligned}
\psi'(t) &= \left(\frac{r\sigma e^{r(\Delta+t)} - 2r\sigma(1-\sigma t)}{\sigma + r(1-\sigma t)}\right) + r\sigma\left(\frac{\sigma e^{r(\Delta+t)} + r(1-\sigma t)^2}{(\sigma + r(1-\sigma t))^2}\right)\\
&= \left(\frac{r\sigma}{[\sigma + r(1-\sigma t)]^2}\right)\left(e^{r(\Delta+t)} - (1-\sigma t)\right)\left(2\sigma + r(1-\sigma t)\right).
\end{aligned}$$

Using the chain of inequalities

$$e^{r(\Delta+t)} > e^{r\Delta} \geq 1 > 1 - \sigma t \geq 0,$$

we can easily verify that $\psi'(t) > 0$ for any $t \in [0, 1/\sigma]$. We conclude that $\psi(t)$ is strictly monotonic on this interval, as desired. $\square$

Appendix C

Proof of Lemma 3.5.

Consider an SSPE with surplus destruction in which every proposer destroys the share $\sigma\bar{t}$ of the current surplus, and receives the share $\bar{x}$ for herself. Suppose by way of contradiction that $\bar{x} \leq \frac{1}{1+(n-1)e^{-r\Delta}}$. Consider a unilateral deviation from this supposed SSPE by some proposer, say Player i. The deviation consists of choosing the pair $(\tilde{\theta}, \tilde{t})$ defined by

$$\begin{aligned}\tilde{\theta}_j &= e^{-r\Delta}\bar{x}, \ j \in N \setminus \{i\},\\ \tilde{\theta}_i &= 1-(n-1)e^{-r\Delta}\bar{x},\\ \tilde{t} &= 0.\end{aligned}$$

From condition (3.3) in Theorem 3.1, it follows that $\tilde{\theta}$ is unanimously accepted. Thus, the unilateral deviation leads to a payoff of $\tilde{\theta}_i$ for Player i. Since unilateral deviations from the supposed SSPE must not be profitable, it follows that $\bar{x} \geq \tilde{\theta}_i$. Moreover, due to Theorem 3.2, Player i's optimal choice of t is unique. Therefore, we must have $\bar{x} > \tilde{\theta}_i$ with strict inequality. It follows that $\bar{x} > 1-(n-1)e^{-r\Delta}\bar{x}$. Rearranging yields $\bar{x} > \frac{1}{1+(n-1)e^{-r\Delta}}$, contradicting the premise that $\bar{x} \leq \frac{1}{1+(n-1)e^{-r\Delta}}$. □

Proof of Theorem 3.8.

We have shown before that $\breve{t}(r)$ is continuous on the whole domain $r \in (0,\infty)$. We have also argued that there are values $\underline{r}$ and $\overline{r}$ such that $\breve{t}(r) = 0$ for all $r \in (0,\infty) \setminus (\underline{r}, \overline{r})$, and $\breve{t}(r) > 0$ for $r \in (\underline{r}, \overline{r})$. Eqns. (3.11)–(3.12) tell us that for every $r \in (\underline{r}, \overline{r})$, the equality $e^{r(\Delta+\breve{t}(r))}(1-\sigma\breve{t}(r))^{-1} = \left(\frac{r}{\sigma}\right)(1-\sigma\breve{t}(r))(n-1)$ holds, and *both* sides of the equality are equal to the proposer premium. For $r \in (0,\infty) \setminus (\underline{r}, \overline{r})$, we know that the proposer premium is equal to $e^{r\Delta}$. Consider the following two piecewise functions:

$$\rho(r) = \begin{cases} e^{r\Delta} & \text{if } (0, \underline{r}], \\ e^{r(\Delta+\breve{t}(r))}(1-\sigma\breve{t}(r))^{-1} & \text{if } (\underline{r}, \overline{r}), \\ e^{r\Delta} & \text{if } [\overline{r}, \infty). \end{cases}$$

$$\widetilde{\rho}(r) = \begin{cases} e^{r\Delta} & \text{if } (0, \underline{r}], \\ \left(\frac{r}{\sigma}\right)(1-\sigma\breve{t}(r))(n-1) & \text{if } (\underline{r}, \overline{r}), \\ e^{r\Delta} & \text{if } [\overline{r}, \infty). \end{cases}$$

It holds that $\rho(r) = \widetilde{\rho}(r)$ for all $r \in (0,\infty)$, and these function assign to every r the proposer premium $\rho(r)$. We want to show that $\rho(r)$ is strictly monotonically increasing in r on the whole domain $(0,\infty)$.

We show first that $\rho(r)$ is continuous on its whole domain $(0,\infty)$. Due to continuity of $\breve{t}(r)$, it is immediate that $\rho(r)$ is piecewise continuous. We have to show its continuity at the points $\underline{r}$ and $\overline{r}$. We know that $\breve{t}(r)$ is continuous and we have the limits $\lim_{r\downarrow\underline{r}} \breve{t}(r) = 0$ and $\lim_{r\uparrow\overline{r}} \breve{t}(r) = 0$. Now, we see that $\lim_{r\downarrow\underline{r}} e^{r(\Delta+\breve{t}(r))}(1-\sigma\breve{t}(r))^{-1} = e^{\underline{r}\Delta}$ and $\lim_{r\uparrow\overline{r}} e^{r(\Delta+\breve{t}(r))}(1-\sigma\breve{t}(r))^{-1} = e^{\overline{r}\Delta}$, as desired. We see that $\rho(r)$ is indeed continuous at the points $\underline{r}$ and $\overline{r}$.

It is immediate that $\rho(r)$ is strictly monotonically increasing on the intervals $(0,\underline{r})$ and $(\overline{r},\infty)$. It remains to be shown that $\rho(r)$ is strictly monotonically increasing on the interval $(\underline{r},\overline{r})$. Take $r^* \in (\underline{r},\overline{r})$, and consider the derivative $\partial\breve{t}/\partial r$ at the point r^*. Suppose first that this derivative is strictly positive. Then, the term $e^{r(\Delta+\breve{t}(r))}(1-\sigma\breve{t}(r))^{-1}$ is strictly increasing in r at the point r^*. So, the derivative $\partial\rho/\partial r$ is strictly positive at the point r^*. Now consider the case where the derivative $\partial\breve{t}/\partial r$ is non–positive at the point r^*. In that case, the term $\left(\frac{r}{\sigma}\right)(1-\sigma\breve{t}(r))(n-1)$ is strictly increasing in r at the point r^*, and we see that the derivative $\partial\widetilde{\rho}/\partial r$ is strictly positive at the point r^*. Since $\widetilde{\rho} = \rho$, this completes the proof.

$\square$

Appendix D

In the model discussed in the main text of the paper, the surplus is diminished by prolonging the time lapse which occurs after a rejection. One may wonder what would happen in a setting where a longer time lapse between bargaining rounds leads to a private cost for the proposer, rather than to a partial destruction of the surplus. In this appendix, we argue that crucial qualitative results of our analysis would be preserved in such a setting. In particular, the following results could be shown: (i) The equilibrium level of surplus destruction depends on the rate of time preference in a continuous and non–monotonic way. If the rate of time preference is either sufficiently small or sufficiently large, no surplus destruction occurs at all. (ii) The proposer premium is monotonically increasing in the rate of time preference. (iii) Whatever the choice of the model parameters, the proposer receives more than the fixed timing allocation. In what follows, we provide a sketch of how these results could be replicated in the model with a private cost to the proposer.

Indeed, let us introduce the following assumption: If a proposer wants to set the time lapse to $\Delta + t$, she pays a private cost σt. The surplus is of unit size and remains unaffected by the choice of t. The remaining elements of the model in the main text of the paper remain unchanged. As in the model in the main text, players are perfectly homogeneous, up to the fact that Player 1 is the first mover. Hence, one could repeat the line of argument in Section 3 and Appendix A to show that an SSPE is "symmetric" in the following sense:

There is a pair (x^*, t^*) such that, in any subgame, the SSPE choice of the time lapse is t^*, and the SSPE division of the surplus gives x^* to the proposer. Due to the private cost assumption, it follows that the proposer's SSPE payoff is $x^* - \sigma t^*$.[11] A player who rejects the current proposal becomes the next proposer after a time lapse of $\Delta + t^*$ so that his payoff from rejecting is $e^{-r(\Delta+t^*)}(x^* - \sigma t^*)$. Therefore, the proposer has to offer this amount to each responder in order to ensure unanimous agreement. Indeed,

$$x^* = 1 - (n-1)(x^* - \sigma t^*)e^{-r(\Delta+t^*)}. \tag{3.19}$$

Appropriately rearranging this expression, the proposer's payoff can be written as

$$x^* - \sigma t^* = \frac{1 - \sigma t^*}{1 + (n-1)e^{-r(\Delta+t^*)}}. \tag{3.20}$$

Moreover, the SSPE time lapse t^* must solve the following problem:

$$\max_{t \in \mathbb{R}_+} 1 - (n-1)(x^* - \sigma t^*)e^{-r(\Delta+t)} - \sigma t. \tag{3.21}$$

[11] In a subgame following previous disagreements, the current proposer may have incurred private costs in previous rounds. These costs need not be taken into account here since they are considered sunk.

From the first–order condition of this problem, we get

$$e^{r(\Delta+t)} = \left(\frac{r}{\sigma}\right)(n-1)(x^* - \sigma t^*).$$

Substituting from Eqn. (3.20) for $x^* - \sigma t^*$, we conclude that the SSPE time lapse t^* is implicitly given as the solution to the following equation:

$$\frac{e^{r(\Delta+t)} + (n-1)}{r} = \frac{(1-\sigma t)(n-1)}{\sigma}. \tag{3.22}$$

This equation is analogous to Eqn. (3.10) in the main text. Both sides of the equation depend continuously on both r and t. Moreover, the right–hand side is independent of r. The left–hand side is a quasi–convex function of r, which goes to infinity both in the limit as r tends to zero and in the limit as r tends to infinity. Therefore, we can use the same arguments as in Subsection 6.1 to show that t depends in a continuous and non–monotonic way on r. Observe also that Eqn. (3.22) admits a solution $t \in (0, 1/\sigma]$ if and only if $(e^{-r\Delta} + n - 1)/r < (n-1)/\sigma$. This is analogous to the conditions for existence of an efficient SSPE or an SSPE with surplus destruction in Theorem 3.5. Therefore, it is true that equilibrium surplus destruction is zero for sufficiently small and sufficiently large r, while it is strictly positive when r falls within an intermediate range. Comparing Eqns. (3.10) and (3.22), it is easy to see that the equilibrium level of surplus destruction with private costs is lower than in the model in the main text. This is very intuitive: "Rent–seeking" is less attractive to the proposer if he fully internalizes its cost.

Now let us show that the proposer premium is strictly monotonically increasing in r. In an SSPE, Player 1 obtains a payoff of $x^* - \sigma t^*$, while each responding player receives $e^{-r(\Delta+t^*)}(x^* - \sigma t^*)$. Hence, the proposer premium defined as the ratio of the proposer's payoff to that of the responder, is $e^{r(\Delta+t^*)}$. We have argued before that the equilibrium level of t is a continuous and non–monotonic function of r. Analogously to Section 6, let us denote that function by $\breve{t}(r)$. From Eqn. (3.22), it follows that both sides of Eqn. (3.23) below must be equal to the proposer premium:

$$e^{r(\Delta+\breve{t}(r))} = \left(\frac{r}{\sigma}\right)(1-\sigma\breve{t}(r))(n-1) - (n-1). \tag{3.23}$$

On an interval where $\breve{t}(r)$ is monotonically increasing in r, the left–hand side of Eqn. (3.23) is monotonically increasing in r. On an interval where $\breve{t}(r)$ is monotonically decreasing in r, the right–hand side of Eqn. (3.23) is monotonically increasing in r. Indeed, an argument analogous to that in the proof of Theorem 3.8 shows that the proposer premium is monotonically increasing in the rate of time preference. Finally, recall that the proof of Lemma 3.5 is based on considering a deviation under which the proposer chooses $t = 0$. Indeed, this proof also applies in the model with private costs. Hence, in an SSPE with surplus destruction, the proposer receives more than his fixed timing allocation, that is,

$x^* \geq \frac{1}{1+(n-1)e^{-r\Delta}}$.

References

Abreu, D. and F. Gul (2000), Bargaining and reputation, *Econometrica, 68,* 85–117.

Avery, C. and P.B. Zemsky (1994), Money Burning and Multiple Equilibria in Bargaining, *Games and Economic Behavior, 7,* 154–168.

Banks, J.S. and J. Duggan (2000), A Bargaining Model of Collective Choice, *American Political Science Review, 94,* 73-88.

Bloch, F. (1996), Sequential Formation of Coalitions in Games with Externalities and a Fixed Payoff Division, *Games and Economic Behavior, 14,* 90–123.

Binmore, K., Osborne, M.J., and A. Rubinstein (1992), Non–cooperative Models of Bargaining, In: Aumann, R.J. and S. Hart (Eds.), Handbook of Game Theory, Volume 1, Elsevier Science Publishers B.V.

Britz, V., P.J.J. Herings and A. Predtetchinski (2010), Non–cooperative Support for the Asymmetric Nash Bargaining Solution, *Journal of Economic Theory, 145,* 1951–1967.

Britz, V., P.J.J. Herings and A. Predtetchinski (2014), On the Convergence to the Nash Bargaining Solution for Action–dependent Protocols, *Games and Economic Behavior, 86,* 178–183.

Busch, L.A., S. Shi and Q. Wen (1998), Bargaining with Surplus Destruction, *The Canadian Journal of Economics, 31,* 915–932.

Cardona, D. and A. Polanski (2013), Voting Rules and Efficiency in One–dimensional Bargaining Games with Endogenous Protocol, *Social Choice and Welfare, 41,* 217–240.

Chatterjee, K., Dutta, B., Ray, D., and K. Sengupta (1993), A Non–cooperative Theory of Coalitional Bargaining, *Review of Economic Studies, 60,* 463–477.

Cunyat, A. (2004), The Optimal Degree of Commitment in a Negotiation with a Deadline, *Economic Theory, 23,* 455–465.

Ellingsen, T. and T. Miettinen (2008), Commitment and Conflict in Bilateral Bargaining, *American Economic Review, 98,* 1629–1635.

Fernandez, R. and J. Glazer (1991), Striking for a Bargain Between two Completely Informed Agents, *American Economic Review, 81,* 240–252.

Haller, H. (1986), Non–cooperative Bargaining of $n \geq 3$ Players, *Economics Letters, 22,* 11–13.

Haller, H. and S. Holden (1990), A Letter to the Editor on Wage Bargaining, *Journal of Economic Theory, 52,* 232–236.

Hart, S. and A. Mas-Colell (1996), Bargaining and Value, *Econometrica, 64,* 357–380.

Herings, P.J.J., Mechalkin, A., and A. Predtetchinski (2017), A One–period Memory Folk Theorem for Multilateral Bargaining Games, *Games and Economic Behavior, 103,* 185–198.

Herrero, M.J. (1985), A Strategic Bargaining Approach to Market Institutions, Ph.D. thesis, London School of Economics.

Imai, H. and H. Salonen (2012), Bargaining and Rent Seeking, Aboa Centre for Economics Discussion Paper No.80, University of Turku.

Kawamori, T. (2008), A Note on Selection of Proposers in Coalitional Bargaining, *International Journal of Game Theory, 37,* 525–532.

Kawamori, T. (2013), Rejector–proposer Legislative Bargaining with Heterogeneous Time and Risk Preferences, *Social Choice and Welfare, 40,* 27–40.

Kawamori, T. (2014), A Non–cooperative Foundation for the Asymmetric Nash Bargaining Solution, *Journal of Mathematical Economics, 52,* 12–15.

Kultti, K. and H. Vartiainen (2010), Multilateral Non–cooperative Bargaining in a General Utility Space, *International Journal of Game Theory, 39,* 677–689.

Laruelle, A. and F. Valenciano (2008), Noncooperative Foundations of Bargaining Power in Committees and the Shapley–Shubik Index, *Games and Economic Behavior, 63,* 341–353.

Li, D. (2011), The Power to Delay, *Economics Letters, 112,* 155–157.

Manzini, P. (1999), Strategic Bargaining with Destructive Power, *Economics Letters, 65,* 315–322.

Miyakawa, T. (2008), Non–cooperative Foundation of n–person Asymmetric Nash Bargaining Solution, *Journal of Economics of Kwansei Gakuin University, 62,* 1–18.

Muthoo, A. (1992), Revocable Commitment and Sequential Bargaining, *The Economic Journal, 102,* 378–387.

Muthoo, A. (1996), A Bargaining Model Based on the Commitment Tactic, *Journal of Economic Theory, 69,* 134–152.

Miettinen, T. and A. Perea (2015), Commitment in Alternating Offers Bargaining, *Mathematical Social Sciences, 76,* 12–18.

Ray, D. (2007), A Game–theoretic Perspective on Coalition Formation, Oxford University Press, ISBN-13: 9780199207954, DOI:10.1093/acprof:oso/9780199207954.001.0001.

Ray, D. and R. Vohra (1999), A Theory of Endogenous Coalition Structures, *Games and Economic Behavior, 26,* 286–336.

Rubinstein, A. (1982), Perfect Equilibrium in a Bargaining Model, *Econometrica, 50,* 97–110.

Schelling, T.C. (1956), An Essay on Bargaining, *American Economic Review, 46,* 281–306.

Schelling, T.C. (1960), The Strategy of Conflict, Cambridge: Harvard University Press.

Selten, R. (1981), A Noncooperative Model of Characteristic Function Bargaining. In: Böhm, V. and H.H. Nachtkamp (Eds.), Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern. Bibliografisches Institut Mannheim, pp. 131–151.

Yildirim, H. (2007), Proposal Power and Majority Rule in Multilateral Bargaining with Costly Recognition, *Journal of Economic Theory, 136,* 167–196.

Yildirim, H. (2010), Distribution of Surplus in Sequential Bargaining with Endogenous Recognition, *Public Choice, 142,* 41–57.

Chapter 4

Negotiating with Frictions

Working Paper

Abstract: We consider bilateral non–cooperative bargaining on the division of a surplus. Compared to the canonical bargaining game in the tradition of Rubinstein, we introduce additional sources of friction into the bargaining process: Implementation of an agreement and consumption of the surplus can only begin at discrete points in time, such as the first day of a month, quarter, or year. Bargaining rounds are of non–trivial length, so that counter–offers may be made without triggering costly delay. Communication between players is noisy: When players make offers, they are uncertain about the time it takes for the offer to arrive. We analyze delays and payoffs in the unique stationary equilibrium of the game. Frictions tend to make the bargaining process less efficient, but lead to a fairer surplus allocation. We establish conditions under which the equilibrium outcome converges to that in a canonical bargaining model as frictions become small.

JEL classification: C72, C78
Keywords: Bargaining, Discount Factor, Timing, Subgame–Perfect Equilibrium, Equilibrium Delay

4.1 Introduction

Many situations in economics, management, and political science can be thought of as surplus division problems: Two or more players own or are able to generate some surplus. However, they can only consume it once they have reached an agreement on its division. One classical economic example is related to labor disputes: When no labor agreement is reached, workers may go on strike, or the firm may lock them out, so that no production takes place. Workers' potential productivity can be thought of as a surplus that is available for consumption if and only if a firm and a union find agreement on its division.

Game theorists have often studied such surplus division problems using non–cooperative bargaining games in the tradition of Rubinstein (1982). This approach considers the bargaining process as a sequence of rounds. In each round, one particular player acts as the proposer and offers some division of the surplus. This offer is then accepted or rejected by the players. Acceptance of an offer ends the game. If an offer is rejected, the game moves to the next round, and the surplus shrinks due to discounting. Throughout the paper, we refer to this setup as the *canonical bargaining model.* One crucial feature of the canonical bargaining model is that offers are made, accepted, and implemented "instantaneously." These steps are all condensed into a single point in time. Thus, the canonical bargaining model assumes that it does not take the proposer any time to make and communicate an offer, nor does the responder need time to evaluate and accept the offer. Moreover, an agreement can be implemented immediately. A costly delay is triggered by a disagreement. One concise way to put it is as follows: The canonical bargaining model assumes that time elapses only between (rather than within) rounds, and a round is defined as starting when a new counter–offer is made.

In the present paper, we take an alternative view. We propose a bilateral bargaining model in which negotiations are subject to frictions that lead to some delay between the points in time when an offer is made, accepted, and implemented. Moreover, bargaining rounds are not defined by a new counter–offer, but by an exogenously fixed time schedule. More specifically, we make three main assumptions:

- Implementation of an agreement and consumption of the surplus can only begin at some exogenously fixed discrete points in time, such as the beginning of each new month, quarter, or year. Delay is only costly when one of these points in time goes by without agreement.

- The time between two potential dates of implementation is what we consider a "bargaining round." During any given bargaining round, one player has the right to make an initial offer at a time of his choosing. The other player may get a chance to make a counter–offer within the same bargaining round.

- When a player makes an offer (or a counter–offer), it is uncertain how long it will take for the opponent to receive it. Bargaining rounds may fail not only if an offer is rejected but also because communication is unsuccessful.

The first assumption drives a wedge between the time at which an offer is accepted and the time at which it can be implemented. Players may find agreement quickly, but have to wait for the next opportunity to implement it. The second assumption qualifies the proposer's privilege that is the driving force behind many results on the canonical bargaining model. Those findings are driven by the idea that each bargaining round is condensed into a single point in time. Under that modeling assumption, it is not meaningful to have a proposer choose the timing of his offer, or to allow a responder to make a counter–offer. In our setup, these considerations become important because each bargaining round is of non–trivial length. One question that we will address is in what sense the findings of the canonical bargaining model can be recovered in the limit as the length of each bargaining round becomes small. The third assumption introduces noise into players' communication, thus creating some friction between the time when an offer is made, and the time when it may be accepted.

In our bargaining model, a proposer faces the following trade–off: On the one hand, if the proposer makes an offer too late, he takes the risk that this offer cannot be implemented because the opponent does not receive it before the envisioned date of its implementation. On the other hand, if the proposer makes an offer way ahead of the date of implementation, the opponent may find it optimal to reject the offer and respond with a counter–offer.

One stylized example of such a bargaining process could be as follows: Suppose that a company and a prospective employee bargain under the institutional or legal constraint that working contracts can only start on the first day of a month. In order to make the hiring decision effective on February 1st, an agreement must be reached in January. Agreeing on January 20th rather than on January 10th has no immediate cost to either party. However, the closer the parties get to January 31st, the more likely it becomes that the negotiation fails due to delays in communication. If no agreement is reached by January 31st, players can continue bargaining in February. However, an agreement reached in February can only be implemented as of March 1st. This delay in implementation is costly.

Restrictions, conventions, or customs relative to the possible dates of implementation of an agreement are common: For instance, in the job market for school teachers or university lecturers, hires typically do not make sense unless they coincide with the start of a semester or academic year. Non–academic jobs typically start with the beginning of a month. Changes to the government's system of taxes or subsidies come into effect with a new fiscal year. More general legislative changes typically become effective with a new month, quarter, or year.

Our main results can be summarized as follows: In the limit as frictions between offer, acceptance, and implementation vanish, efficiency is maximized, and the ex ante expected division of the surplus corresponds to that familiar from the canonical bargaining model. If communication among players involves substantial friction, however, proposals remain lopsided even if offers are arbitrarily frequent. However, the relative payoffs of players need not depart from the canonical predictions. If potential dates of implementation are sufficiently far apart, equilibrium outcomes tend to be more fair but less efficient.

The present paper is related to two main strands of the non–cooperative bargaining literature. One strand looks at the canonical bargaining model with a variety of proposer selection protocols, and examines how exactly the protocol determines the equilibrium division of the surplus. This relationship between the distribution of proposal power and the equilibrium allocation has been explored in great detail by Hart and MasColell (1996), Laruelle and Valenciano (2008), Miyakawa (2008), Kultti and Vartiainen (2010), and Britz et al. (2010), among others.

Some related papers have studied non–cooperative bargaining in the presence of a deadline, and have suggested explanations for *deadline effects,* that is, agreements tend to be reached close to the deadline after some period of delay, see for instance Fershtman and Seidmann (1993) or Ponsati (1995). Ma and Manove (1993) consider a model in which two players negotiate in the presence of a deadline and communication is noisy.

The remainder of this paper is organized as follows: The formal model description is presented in Section 2. Then, Section 3 contains the analysis of SSPE of this model. Section 4 is devoted to some comparative statics analyses. Section 5 investigates the socially optimal length of bargaining rounds from an efficiency and fairness point of view. Section 6 concludes.

4.2 Bargaining game

Two players decide on the division of a surplus by non–cooperative bargaining. While bargaining takes place in continuous time $[0, \infty)$, an agreement can only be implemented at equidistant points in time $T,\ 2T,\ 3T, \ldots$, where $T > 0$ is exogenously fixed. We refer to the time interval $[0, T]$ as the first *bargaining round,* and more generally to the time interval $[(k-1)T,\ kT]$ as the k^{th} bargaining round, for $k = 1, 2, \ldots$.

Players have a common rate of time preference $r > 0$. When it is convenient, we will work with the discount factor $\delta = e^{-rT}$ instead. As a normalization, we assume that the surplus is of size one. This implies that the discounted value of an agreement reached in the first bargaining round is δ.[1] More generally, the discounted value of an agreement in

[1] It might seem more intuitive to assume that the physical surplus is of size $e^{rT} = 1/\delta$, so that its value is one if an agreement is implemented at the earliest possible moment. Later in the paper, however, we are going to vary T independently of r in a comparative statics analysis. Thus, we have to fix the size of the physical surplus to one.

bargaining round k is δ^k.

In each bargaining round, one player is the *proposer* and the other player is the *responder.* Each bargaining round k proceeds as follows:

The proposer offers some split of the surplus, to become effective at time kT. He is free to choose at what time during bargaining round k he makes this offer. More formally, the proposer chooses a pair $(\theta, \tau) \in [0,1] \times [0,T]$, where θ indicates the amount of surplus which the proposer offers to the responder, while $(k-1)T + \tau$ is the time at which he makes this offer. A noisy channel of communication is then used to transmit the offer to the responder. More specifically, we model communication as a Poisson process with arrival rate $\lambda > 0$. This implies that an offer made at time $(k-1)T + \tau$ reaches the responder before time kT with probability $1 - e^{-\lambda(T-\tau)}$.

If the offer fails to arrive until time kT, then we say that bargaining round k *fails,* and the game moves to round $k+1$. Note that this setup allows the proposer to effectively pass the opportunity to propose: He can do so by delaying his offer until time kT. We will see, however, that it is never optimal for the proposer to pass. If the responder receives the proposer's offer before time kT, the responder can either accept it or make a counter–offer. If the proposer's offer is accepted, the game ends and the proposer and responder receive utilities $\delta^k(1-\theta)$ and $\delta^k\,\theta$, respectively.[2] Now suppose that the responder does not accept the offer but makes a counter–offer $\eta \in [0,1]$. Again, an uncertain amount of time elapses until the responder's counter–offer reaches the proposer. As before, this delay in communication is modeled by a Poisson process with arrival rate λ. [3] If the counter–offer does not arrive until time kT, then round k ends in disagreement, and the game moves to round $k+1$. If the responder's counter–offer does arrive before time kT, the proposer chooses to accept or reject it. If he rejects it, then round k ends in disagreement, and the game moves to round $k+1$. If he accepts, then the game ends and the proposer and responder receive payoffs $\delta^k\,\eta$ and $\delta^k(1-\eta)$, respectively. Whenever an agreement is not reached by time kT, we say that bargaining round k *fails.*

It remains to specify how the proposer in each round is chosen: Without loss of generality, we assume that Player 1 is the proposer in the first bargaining round. Moreover, for any $k \geq 2$, we assume that if Player $i = 1,2$ is the proposer in bargaining round $k-1$, then Player i is also the proposer in bargaining round k with probability m_i. With complementary probability $1 - m_i$, Player $j \neq i$ is the proposer in round k. Hence, the

[2] Note that, if an offer is accepted, the proposer receives the complement of what he offered to the responder. This amounts to a tacit assumption that offers must be efficient. We could have allowed the proposer to make an offer which leaves some surplus unallocated. In equilibrium, such an offer would not be made, so our results would not change.

[3] We assume there that if the responder makes a counter–offer she does so immediately after receiving the proposer's offer. We could have assumed instead that the responder also has the possibility to wait before making a counter–offer. This would not change the results, however: We will see that, in equilibrium, the responder has no incentive to wait.

proposer selection follows a Markov chain with the transition matrix

$$M = \begin{pmatrix} m_1 & 1 - m_1 \\ 1 - m_2 & m_2 \end{pmatrix}.$$

We assume that $m_i < 1$ for each $i = 1, 2$, so that the Markov chain is irreducible. Its stationary distribution $\mu = (\mu_1, \mu_2)$ is given by $\mu M = \mu$, and can be written as

$$\mu_i = (1 - m_j)/(2 - m_i - m_j),$$

for each $i = 1, 2$ and $j \neq i$. One noteworthy special case is $m_1 = m_2 = 0$, which means that the role of proposer alternates from one round to the next. This is the proposer selection protocol in Rubinstein's original paper.

A *stationary strategy* for Player i consists of the following elements:

- A pair (θ_i, τ_i) such that if Player i is the proposer in round k, he proposes θ_i at time $kT + \tau_i$.
- A map $\alpha_i : [0,1] \times [0,T] \to \{\mathit{Accepts}\} \cup [0,1]$ such that in any round k in which Player i is the responder, he accepts an offer θ which reaches him at time $kT + t$ if and only if $\alpha_i(\theta, t) = \mathit{Accept}$, and otherwise, he makes the counter–offer given by $\alpha_i(\theta, t)$.
- A map $\beta_i : [0,1] \times [0,T] \to \{\mathit{Accept}, \mathit{Reject}\}$ such that in any round k in which Player i is the proposer, he accepts a counter–offer η which reaches him at time $kT + t$ if and only if $\beta_i(\eta, t) = \mathit{Accept}$, and otherwise, he rejects.

A *stationary subgame–perfect equilibrium (SSPE)* is a profile of stationary strategies which is a subgame–perfect Nash equilibrium.

4.3 Analysis of stationary equilibrium

The purpose of this section is to derive expressions for the expected payoffs and expected delay in an SSPE. The starting point for this analysis is the stationarity property of the game just described: The subgame starting at time $(k-1)T$ for any $k \geq 2$ is equivalent to the entire game, up to the identity of the proposer. Thus, in any SSPE, there exists a quadruple (x_1, x_2, y_1, y_2) such that if Player $i = 1, 2$ is the proposer in round k, then the expected SSPE payoffs for Player i and Player $j \neq i$ in the subgame starting at time $(k-1)T$ are $\delta^k x_i$ and $\delta^k y_j$, respectively.

In what follows, we will consider a bargaining round k in which Player $i = 1, 2$ is the proposer and Player $j \neq i$ is the responder. To this end, it is useful to define the following auxiliary variables:

$$\begin{aligned}\widetilde{x}_i &= m_i x_i + (1-m_i)y_i,\\ \widetilde{y}_j &= m_i y_j + (1-m_i)x_j.\end{aligned}$$

Player i continues to act as proposer in round $k+1$ with probability m_i. Hence, $\delta\widetilde{x}_i$ and $\delta\widetilde{y}_j$ are reservation payoffs for Players i and j, respectively.

Recall that the rejection of a counter–offer by the proposer implies that bargaining can resume only in the next round. Hence, it is straightforward that in an SSPE, Player i accepts Player j's counter–offer η_j if and only if $\eta_j \geq \delta\widetilde{x}_i$. Note that this is independent of the time at which Player i receives the counter–offer. Therefore, it would never be optimal for Player j to wait before making a counter–offer. This justifies our simplifying assumption that counter–offers are made without delay.

Now we proceed by backward induction to a history where Player j decides whether to accept Player i's offer θ_i, or to make a counter–offer. Intuitively, the later Player j receives an offer θ_i, the more risky it is for him to send a counter–offer. During each bargaining round, the responder's bargaining position gradually erodes overtime. As time approaches kT, Player j becomes more willing to make concessions to Player i. We will show that, for any given offer by Player i, there is some critical time from which onwards Player j accepts the offer. The proposition below claims that this critical point in time is given by

$$\widehat{t}_j(\theta_i) = \begin{cases} \max\left\{0,\ T - \left(\frac{1}{\lambda}\right)\ln\left(\frac{1-\delta\widetilde{x}_i-\delta\widetilde{y}_j}{1-\delta\widetilde{x}_i-\theta_i}\right)\right\} & \text{if } \theta_i < 1-\delta\widetilde{x}_i,\\ 0 & \text{if } \theta_i \geq 1-\delta\widetilde{x}_i.\end{cases}$$

The formal proof of the proposition is relegated to Appendix A.

Proposition 4.1. *Suppose that Player j receives Player i's offer θ_i at time $(k-1)T+t$. In an SSPE, Player j accepts if and only if $t \geq \widehat{t}_j(\theta_i)$.*

Observe that

$$\lim_{\theta_i \uparrow 1-\delta\widetilde{x}_i} T - \left(\frac{1}{\lambda}\right)\ln\left(\frac{1-\delta\widetilde{x}_i-\delta\widetilde{y}_j}{1-\delta\widetilde{x}_i-\theta_i}\right) = -\infty.$$

Hence, there is $\varepsilon > 0$ sufficiently small so that Player j accepts proposal $\theta_i = 1-\delta\widetilde{x}_i-\varepsilon$ at any time during round k. As a result, the proposer can always obtain an expected payoff strictly greater than his reservation level $\delta\widetilde{x}_i$. In particular, a proposer never finds it optimal to "pass" his opportunity to make an offer by waiting until kT. This is formally stated in the proposition below. The proof can be found in Appendix A.

Proposition 4.2. *In an SSPE, Player i makes an offer θ_i such that $\theta_i < 1-\delta\widetilde{x}_i$. Moreover, he makes the offer strictly earlier than at time kT.*

The next step is to show that it is optimal for Player i to make his offer exactly at the critical point in time when Player j is ready to accept it. The intuition is as follows: Player j is ready to accept a given offer from some critical point in time onwards. On the one hand, if Player i makes the offer before Player j is ready to accept it, there is a risk that the offer arrives so soon that Player j will prefer to make a counter–offer. On the other hand, if Player i makes the offer when Player j is already willing to accept it, then Player i could improve the probability of acceptance by making the same offer sightly earlier. One implication of the next proposition is that, on the path of play of an SSPE, no counter–offers will ever be made.

Proposition 4.3. *In an SSPE, Player i chooses the pair (θ_i, τ_i) in such a way that $\tau_i = \widehat{t}_j(\theta_i)$ and $\theta_i = (1 - e^{-\lambda(T-\tau_i)})(1 - \delta\widetilde{x}_i) + e^{-\lambda(T-\tau_i)}\delta\widetilde{y}_j$.*

The proof of Proposition 4.3 is relegated to Appendix A.

We have shown that no counter–offers are made on an equilibrium path of play. Hence, the only way how bargaining round k can fail is if Player i's offer does not arrive before time kT. For any $\tau_i \in [0, T]$, let us call $\pi_i(\tau_i) = e^{-\lambda(T-\tau_i)}$ the *failure probability* of an offer made at time $kT + \tau_i$. Throughout the paper, we will omit the argument τ_i if no confusion arises. One implication of Proposition 4.3 is that Player i's optimal choice of a pair (θ_i, τ_i) can be reduced to an optimal choice of the failure probability π_i. In case of failure, Player i expects to get $\delta\widetilde{x}_i$ in the ensuing continuation game. If bargaining round k does not fail, then Player i receives $1 - \theta_i > \delta\widetilde{x}_i$, where θ_i is implicitly determined by the choice of π_i in the way specified by Proposition 4.3. Therefore, Player i's expected payoff in round k equals $\delta^k\xi_i(\pi_i)$, where $\xi_i(\pi_i)$ is given by

$$\xi_i(\pi_i) = \pi_i\delta\widetilde{x}_i + (1 - \pi_i)(1 - (1 - \pi_i)(1 - \delta\widetilde{x}_i) - \pi_i\delta\widetilde{y}_j).$$

After some simplification, we obtain

$$\xi_i(\pi_i) = \delta\widetilde{x}_i + (1 - \delta\widetilde{x}_i - \delta\widetilde{y}_j)(\pi_i - \pi_i^2). \tag{4.1}$$

This expression for Player i's expected payoff has a straightforward interpretation: The first summand $\delta\widetilde{x}_i$ is the (discounted) continuation payoff for Player i in the next bargaining round. The expression $1 - \delta\widetilde{x}_i - \delta\widetilde{y}_j$ represents the share of the surplus which the players forgo if the current round fails. Put another way, we can interpret it as the *gain from immediate agreement.* This gain realizes with probability $1 - \pi_i$, and if it does, then our analysis so far reveals that Player i gets a share π_i of it. On the one hand, Player i can trivially ensure that his offer fails with probability one by making it only at the deadline kT. On the other hand, by making the offer immediately at time $(k-1)T$, he can reduce the failure probability in round k to $e^{-\lambda T}$. More formally, Player i's optimization problem can be written as

$$\max_{\pi_i} \ \xi_i(\pi_i) \text{ subject to } \pi_i \in [e^{-\lambda T}, 1]. \tag{4.2}$$

Consider the derivative

$$\partial\xi_i/\partial\pi_i = (1 - \delta\widetilde{x}_i - \delta\widetilde{y}_j)(1 - 2\pi_i).$$

We observe that $\widetilde{x}_i + \widetilde{y}_j \leq 1$ and so $1 - \delta\widetilde{x}_i - \delta\widetilde{y}_j \geq 1 - \delta > 0$. There are two cases to distinguish: Suppose first that $e^{-\lambda T} \leq 1/2$. In this case, the first–order condition

$$\partial\xi_i/\partial\pi_i = 0$$

yields $\pi_i = 1/2$. Now suppose that $e^{-\lambda T} > 1/2$. In that case, $\partial\xi_i/\partial\pi_i < 0$ for any $\pi_i \in [e^{-\lambda T}, 1]$, and hence Player i finds it optimal to choose $\pi_i = e^{-\lambda T}$. We note that the exact values of $\widetilde{x}_i$ and $\widetilde{y}_j$, as well as the identities of Players i and j, do not affect the solution to this optimization problem. Hence, we have the following proposition:

Proposition 4.4. *In an SSPE, τ_1 and τ_2 are chosen such that* $\pi_1 = \pi_2 = \max\left\{\frac{1}{2}, e^{-\lambda T}\right\}$.

It is noteworthy that this result is very general: It does not depend on the continuation utilities that players expect from the next round. In particular, it is independent of players' time preferences. This is somewhat surprising: One might expect that players are more willing to risk bargaining failure, and hence a delay, when they are more patient. It turns out that this is not true.

Moreover, if T is sufficiently large, the gain from immediate agreement is split fairly between the two players. There is no proposer premium. This is a consequence of our assumption that players can respond to offers by counter–offers without triggering a costly delay. Nevertheless, the complete absence of a proposer premium is not trivial: When a counter–offer is made, the potential date of implementation is closer, and thus the risk of costly delay greater than when an initial offer is made. One may therefore have intuitively expected that, even in our model, the bargaining position of the proposer is always stronger than that of the responder.

Proposition 4.4 has a number of additional implications that will be important in the remainder of this paper. One of them is the following corollary:

Corollary 4.1. *The proposer's equilibrium offer is given by*

$$\theta_i = \begin{cases} \frac{1}{2}(1 - \delta\widetilde{x}_i) + \frac{1}{2}\delta\widetilde{y}_j & \text{if } e^{-\lambda T} \leq \frac{1}{2}, \\ (1 - e^{-\lambda T})(1 - \delta\widetilde{x}_i) + e^{-\lambda T}\delta\widetilde{y}_j & \text{if } e^{-\lambda T} \geq \frac{1}{2}. \end{cases} \tag{4.3}$$

Proposition 4.4 says that, in an SSPE, every bargaining round which is reached fails with equal probability $\pi = \max\left\{\frac{1}{2}, e^{-\lambda T}\right\}$. This allows us to write the expected size of the surplus at the time of an agreement, discounted back to the beginning of the game, as

$$\delta v(\pi) = \delta(1-\pi)\sum_{k=0}^{\infty}(\delta\pi)^k = \delta\left(\frac{1-\pi}{1-\delta\pi}\right).$$

Note that the amount $\delta v(\pi)$ is independent of the initial proposer's identity. Therefore,

$$x_1 + y_2 = x_2 + y_1 = v(\pi). \tag{4.4}$$

In what follows, we will refer to the quantity $\delta v(\pi)$ as the *expected value of agreement.* In a similar manner, we can also compute the expected time at which an agreement is implemented as

$$\begin{aligned} \omega(\pi, T) = &\ T(1-\pi) + 2T\pi(1-\pi) + 3T\pi^2(1-\pi) + \dots \\ = &\ T(1-\pi)\textstyle\sum_{k=0}^{\infty}(k+1)\pi^k \\ = &\ T/(1-\pi). \end{aligned}$$

From now on, we will refer to this quantity as the *expected implementation time.*

In equilibrium, the failure probability π is chosen optimally, so that the expected value of agreement and the expected waiting time can also be thought of as the following functions of the model parameters $r,\ \lambda$, and T:

$$\omega(\lambda, T) = \begin{cases} 2T & \text{if } T \geq \ln(2)/\lambda, \\ \frac{T}{1-e^{-\lambda T}} & \text{if } T \leq \ln(2)/\lambda. \end{cases}$$

$$v(r, \lambda, T) = \begin{cases} \frac{1}{2e^{rT}-1} & \text{if } T \geq \ln(2)/\lambda, \\ \frac{e^{\lambda T}-1}{e^{(\lambda+r)T}-1} & \text{if } T \leq \ln(2)\lambda. \end{cases}$$

We have now introduced all preliminaries that we need to characterize SSPE payoffs and delay.

Our starting point was the existence of $x_i,\ y_j$ such that if Player i is the proposer in round k, the SSPE utilities in the subgame starting in that round are given by $\delta^k x_i$ and $\delta^k y_j$. By definition of ξ_i and by Proposition 4.4, it follows that

$$\xi_i\left(\max\left\{\frac{1}{2}, e^{-\lambda T}\right\}\right) = x_i.$$

We can now conclude that, for any given δ, SSPE utilities x_1, x_2, y_1, and y_2 and associated continuation utilities $\widetilde{x}_1, \widetilde{x}_2, \widetilde{y}_1$, and $\widetilde{y}_2$ solve the following system of equations:

$$\begin{aligned}
x_1 &= \delta\widetilde{x}_1 + (1-\delta v)(\pi - \pi^2), && (4.5)\\
x_2 &= \delta\widetilde{x}_2 + (1-\delta v)(\pi - \pi^2), && (4.6)\\
y_1 &= v - x_2, && (4.7)\\
y_2 &= v - x_1, && (4.8)\\
\widetilde{x}_1 &= m_1 x_1 + (1-m_1) y_1, && (4.9)\\
\widetilde{x}_2 &= m_2 x_2 + (1-m_2) y_2, && (4.10)\\
\widetilde{y}_1 &= m_2 y_1 + (1-m_2) x_1, && (4.11)\\
\widetilde{y}_2 &= m_1 y_2 + (1-m_1) x_2, && (4.12)
\end{aligned}$$

where

$$v = (1-\pi)/(1-\delta\pi).$$

Eqns. (4.5)–(4.12) are a system of eight linearly independent equations in eight unknowns: A solution exists, and it is unique. For the remainder of our discussion, the solutions for the variables x_1 and y_2 are particularly important.

We have assumed that Player 1 is the initial proposer, so the SSPE payoffs in the entire game are δx_1 and δy_2, where the quantities x_1 and y_2 are given by

$$\begin{aligned}
x_1 &= \left(\frac{1-\pi}{1-\delta\pi}\right)\left(\frac{\delta + (1-\delta)\pi - \delta m_1 + \delta\pi(m_1 - m_2)}{1-\delta(m_1+m_2-1)}\right), && (4.13)\\
y_2 &= \left(\frac{1-\pi}{1-\delta\pi}\right)\left(\frac{1 - (1-\delta)\pi - \delta m_2 - \delta\pi(m_1 - m_2)}{1-\delta(m_1+m_2-1)}\right). && (4.14)
\end{aligned}$$

In Appendix B, we provide Mathematica code that can be used to verify our computation of the solution to this system of equations.

Let the tuple

$$(x_1^*, x_2^*, y_1^*, y_2^*, \widetilde{x}_1^*, \widetilde{x}_2^*, \widetilde{y}_1^*, \widetilde{y}_2^*)$$

solve the system of Eqns. (4.5)–(4.12). Define a profile of stationary strategies

$$\sigma^* = (\theta_1^*, \theta_2^*, \tau_1^*, \tau_2^*, \alpha_1^*, \alpha_2^*, \beta_1^*, \beta_2^*)$$

as follows:

- For each $i = 1, 2$, let $\theta_i^* = (1-\pi)(1-\delta\widetilde{x}_i^*) + \pi\delta\widetilde{y}_j^*$, where $\pi = \max\{\frac{1}{2}, e^{-\lambda T}\}$.
- For each $i = 1, 2$, let $\tau_i^* = \widehat{t}_j(\theta_i^*)$.
- For each $j = 1, 2$, let $\alpha_j^*(\theta, t) = \textit{Accept}$ if and only if
  $\theta \geq (1 - e^{-\lambda(T-t)})(1-\delta\widetilde{x}_i^*) + e^{-\lambda(T-t)}\delta\widetilde{y}_j^*$.
  Otherwise, $\alpha_j^*(\theta, t) = \delta\widetilde{x}_i^*$.

- Let $\beta_i^*(\eta, t) = \mathit{Accept}$ if and only if $\eta \geq \delta \widetilde{x}_i^*$, and let $\beta_i^*(\eta, t) = \mathit{Reject}$ otherwise.

Theorem 4.1. *The profile of stationary strategies σ^* is the unique SSPE.*

Proof. Uniqueness of SSPE follows from the fact that, in an SSPE, the utilities and associated continuation utilities must satisfy Eqns. (4.5)–(4.12), and that system of equations has a unique solution. In order to show that the strategy profile σ^* is indeed an SSPE, it is straightforward that the accept/reject behavior prescribed by α_i^* and β_i^* is optimal. Moreover, the optimality of (θ_i^*, τ_i^*) follows by applying the same logic as in the proof of Proposition 4.3, and then observing that the solution to the optimization problem (4.2) is unique. □

We refer to the ratio $\rho_1 = x_1/y_2$ as Player 1's *relative payoff;* it is given by

$$\rho_1 = \frac{\delta + (1-\delta)\pi - \delta m_1 + \delta\pi(m_1 - m_2)}{1 - (1-\delta)\pi - \delta m_2 - \delta\pi(m_1 - m_2)}. \tag{4.15}$$

In what follows, we explore how equilibrium variables depend on the underlying model parameters. Some relevant observations can be made immediately from the above expressions:

Consider the case where $T > \ln(2)/\lambda$ and therefore the equilibrium choice of the failure probability is $\pi = 1/2$. In that case, we observe that

$$x_1 = y_2 = \left(\frac{1}{2-\delta}\right)\left(\frac{\frac{1}{2} + \frac{\delta}{2}(m_1 + m_2 - 1)}{1 - \delta(m_1 + m_2 - 1)}\right) = \left(\frac{1}{2-\delta}\right)\left(\frac{1}{2}\right).$$

Thus, the relative payoff ρ_1 equals one regardless of the value of the discount factor and the transition probabilities: For any δ and any m_1 and m_2, Players 1 and 2 receive the same expected SSPE payoff. In the canonical Rubinstein bargaining game, the discount factor and transition probabilities are the crucial determinants of bargaining power. This is not true in our setting when $T \geq \ln(2)/\lambda$: In that case, there is no advantage from being the proposer and consequently, the equilibrium surplus allocation does not depend on the proposer selection protocol. Along any equilibrium path, the proposer always offers the fair split, and the responder accepts. However, there is some degree of inefficiency which comes from the fact that each bargaining round which is reached on the equilibrium path fails with probability 1/2. Unlike in the canonical bargaining model, the expected value of agreement does not equal one but rather

$$\delta v = \delta/(2-\delta),$$

and so the expected payoff to each player is $\delta/(4-2\delta)$. This converges to 1/2 in the limit as $\delta \to 1$. In our model, when the bargaining rounds are of sufficient length, an increase in the discount factor boosts efficiency but does not change the surplus distribution. This is exactly the reverse as in the canonical bargaining model, where an increase in the discount factor makes the payoff distribution more fair, while the bargaining outcome is efficient regardless of the discount factor.

Now suppose that bargaining rounds are sufficiently short so that $e^{-\lambda T} = \pi > 1/2$. In that case, there is a proposer advantage like in the canonical bargaining model. The proposer's share does depend on the discount factor. Here, a "proposer advantage" means that the proposer gets more than half of the gain from immediate agreement in each round. In the canonical bargaining model, he gets the whole gain from immediate agreement in each round. Like in the canonical bargaining model, relative payoffs depend on the proposer protocol. In particular, we find

$$\lim_{\delta \to 1} \rho_1 = \left(\frac{1 - m_1(1-\pi) - \pi m_2}{1 - m_2(1-\pi) - \pi m_1} \right) = \left(\frac{1 - m_1 + \pi(m_1 - m_2)}{1 - m_2 - \pi(m_1 - m_2)} \right).$$

The impact of proposal power on the equilibrium allocation is weaker than in the canonical bargaining model. This is because the non–trivial length of bargaining rounds gives the responder a chance to make a counter–offer. This attenuates the extent of the proposer's strategic advantage over the responder.

Example 4.1. *Let us consider the numerical example with* $\lambda = 1$ *and* $T = 0.5$. *Since* $T = 0.5 < \ln(2) \approx 0.69$, *we have that* $\pi = e^{-0.5} \approx 0.607$. *Moreover, let us suppose that* $m_1 = 0.6$ *and* $m_2 = 0.4$. *This implies that the probabilities in the stationary distribution of the Markov chain are also* $\mu_1 = 0.6$ *and* $\mu_2 = 0.4$. *The limit of relative payoffs is*

$$\lim_{\delta \to 1} \rho_1 = \frac{0.4 + 0.2e^{-0.5}}{0.6 - 0.2e^{-0.5}} \approx 1.089.$$

In a canonical bargaining game in which proposer selection follows the same transition dynamics, one would expect the limit of relative payoff of Player 1 to be $\mu_1/\mu_2 = 0.6/0.4 = 1.5$.

4.4 The degree of noise

We are now going to examine how SSPE payoffs and delays change as we vary the arrival rate λ. First suppose that λ is arbitrarily large. In that case, Proposition 4.4 tells us that the proposer in bargaining round k waits almost until the deadline kT, thus ensuring that the proposal fails to arrive before kT with probability one half. The expected time for implementation of the agreement and the expected equilibrium payoffs remain constant once λ has grown beyond the point where $e^{-\lambda T} \geq 1/2$.

Proposition 4.5. *For any given r and T, in the limit as λ is sufficiently large, every proposer along an SSPE path of play offers the fair split, every bargaining round fails with probability $1/2$, and expected implementation time is $2T$.*

As $\lambda \to \infty$, the friction arising from noisy communication vanishes. This does not make the bargaining process more efficient, however: The proposer adjusts the timing of his offer in a way that holds the failure probability constant at one half. No matter how close to instantaneous the communication between the players is, there remains a substantial expected delay in equilibrium.

Now we turn to the case where λ is small. As a benchmark, let us briefly consider a canonical bargaining model with two players and linear utilities in which proposer selection follows a Markov chain.[4] Let $\widehat{x}_i$ and $\widehat{y}_j$ be the SSPE utilities of a proposer and a responder in any subgame of that canonical bargaining model. It is well–known that these utilities are given by the solution to the following system of equations:

$$\begin{aligned} \widehat{x}_i &= 1 - \widehat{y}_j, \\ \widehat{y}_j &= \delta m_i \widehat{y}_j + \delta(1 - m_i)\widehat{x}_j, \end{aligned}$$

for $i = 1, 2$ and $j \neq i$. The former equation follows from the fact that an SSPE of the canonical bargaining model is always efficient. In particular, agreement is reached immediately in every subgame. The latter equation captures the standard result that in an SSPE, the responder is indifferent between acceptance and rejection of the proposal. Solving this system yields

$$\begin{aligned} \widehat{x}_i &= \frac{1 - \delta m_j}{1 - \delta(m_i + m_j - 1)}, \\ \widehat{y}_j &= \frac{\delta - \delta m_i}{1 - \delta(m_i + m_j - 1)}, \end{aligned}$$

for $i = 1, 2$ and $j \neq i$. The *relative payoffs* in an SSPE of the canonical bargaining model are given by the ratio

$$\widehat{\rho}_i = \frac{1 - \delta m_j}{\delta - \delta m_i}.$$

We will now use the canonical bargaining model as a benchmark to which we compare the SSPE of our bargaining model in the limit as λ goes to zero while r and T remain fixed. In that case, it is straightforward that the expected implementation time grows

[4]The model and resutls briefly sketched here as a benchmark are a simple special case of Britz et al. (2010): That paper studies a canonical bargaining model where proposer selection follows a Markov chain. The model in that paper is much more general, however, in the sense that it allows for an arbitrary finite number of players and a general convex set of feasible utilities.

without bound, and so the expected value of agreement converges to zero, indeed:

$$\lim_{\lambda \to 0} \delta \left(\frac{1 - e^{-\lambda T}}{1 - \delta e^{-\lambda T}} \right) = 0.$$

When λ is small enough, Player i's offer θ_i is approximately equal to Player j's continuation utility $\delta \widetilde{y}_j$. This continuation utility in turn is bounded above by the expected value of agreement δv, which again is monotone decreasing in λ and converges to zero in the limit as λ goes to zero. From inspection of Eqn. (4.15), we find the limit of relative payoffs

$$\lim_{\lambda \to 0} \rho_1 = \frac{1 - \delta m_2}{\delta - \delta m_1},$$

and so we have the following proposition:

Proposition 4.6. *For any δ, in the limit as $\lambda \to 0$, the expected value of agreement converges to zero. Relative expected payoffs converge to the same limit as those in the canonical bargaining model. Along an equilibrium path of play, however, proposers offer almost nothing to responders.*

Intuitively, if communication is very noisy, this weakens the responder's bargaining position: Once an offer has been received, making a counter–offer would likely lead to a substantial and costly delay. Thus, a proposer only has to offer very little surplus to a responder. In that sense, when communication is noisy enough, bargaining is in a standoff: Each player finds it optimal to insist on almost the entire surplus. As a result, an equilibrium path of play looks as follows: Players keep making very lopsided offers to each other. Each of these offers is unlikely to arrive in time, so that bargaining likely keeps failing for many rounds until eventually some offer arrives in time to be accepted. A player's relative bargaining power depends on the probability with which one of his offers is eventually successful. This probability corresponds to the share of time for which this player expects to be the proposer in the long–run. This makes it intuitively clear why, with small λ, relative payoffs are driven by the distribution of proposal power in the same way as in the canonical bargaining model – even though the realized outcome is much more lopsided, and absolute expected payoffs are small.

4.5 Optimal length of bargaining rounds

One key assumption of the present paper is that there are equidistant points in time T, $2T$, ... at which an agreement can be implemented, and consumption of the surplus can begin. Players make offers on how to split the surplus at the next available date. One important question is how our equilibrium predictions vary with the choice of the institutional parameter T. In particular, one might wonder which value of T is preferred by Player 1 or Player 2, and which one is preferable from an efficiency or fairness point of view. Observe that changing T has two effects: First, the length of the bargaining rounds influences the failure probability. If rounds are short enough, the proposer cannot choose his optimal failure probability of one half, but has to settle for a higher failure probability. Shorter bargaining rounds are more likely to fail. Second, when bargaining rounds are shorter, their failure becomes less costly: The next opportunity to implement agreement is less far away. So far, we have done some comparative statics analysis on the parameters r and λ, allowing us to vary the discount factor δ and the failure probability π independently of each other. As we consider changes in T, however, we are simultaneously varying both the discount factor $\delta = e^{-rT}$ and the failure probability $\pi = \max\{\frac{1}{2}, e^{-\lambda T}\}$. Both of them are non–decreasing in T. They both converge to one in the limit as $T \to 0$.

4.5.1 The limit case

For the purpose of this section, let us consider the expected value of agreement and the expected implementation time as a function of T, while keeping r and λ fixed, thus:

$$v(T) = \begin{cases} \frac{e^{\lambda T}-1}{e^{(\lambda+r)T}-1} & \text{if } T \leq \ln(2)/\lambda, \\ \frac{1}{2e^{rT}-1} & \text{if } T > \ln(2)/\lambda. \end{cases}$$

$$\omega(T) = \begin{cases} \left(\frac{e^{\lambda T}}{e^{\lambda T}-1}\right) T & \text{if } T \leq \ln(2)/\lambda, \\ 2T & \text{if } T > \ln(2)/\lambda. \end{cases}$$

Note that both the expected value of agreement and the expected implementation time are continuous functions of T. In particular, they are continuous at the point $T = \ln(2)/\lambda$. It is readily apparent from the above expressions that:

Proposition 4.7. *The expected value of agreement in an SSPE is strictly monotonically decreasing in the length of bargaining rounds.*

The intuition is as follows: The bargaining process can be thought of as a sequence of three consecutive stages: First, bargaining may be suspended for some time because the proposer waits for the optimal moment to make an offer. Second, an offer has been made by the proposer but not yet received by the responder. Third, an agreement has been reached but the time for its implementation has not come yet. If bargaining rounds are sufficiently short, a proposer never waits before making a counter–offer which eliminates the first stage. Moreover, as bargaining rounds become shorter, an agreement is reached closer to the end of a round, and hence closer to a possible date of implementation. Thus, the third stage also vanishes. If T is small enough, a bargaining round tends to consist entirely of the time that it takes for the offer to be communicated from the proposer to the responder. Hence, it is clear that for small T, the expected implementation time must converge to the expected arrival time of the underlying Poisson process. Indeed, by applying L'Hôpital's rule, we can easily verify that

$$\lim_{T\to 0}\ \omega(T) = \lim_{T\to 0}\left(\frac{e^{\lambda T}}{e^{\lambda T}-1}\right)T = 1/\lambda.$$

Since the expected value of agreement is monotone decreasing in T, it is bounded above by its limit as $T\to 0$. Again applying L'Hôpital's rule, this limit can be computed as

$$\lim_{T\to 0}\delta v(T) = \lim_{T\to 0}\left(\frac{e^{\lambda T}-1}{e^{(\lambda+r)T}-1}\right) = \frac{\lambda}{\lambda+r}.$$

While delays within a round are costless if the length of the rounds is exogenously fixed, they do become an important consideration once we study the optimal length of bargaining rounds.

Finally, we consider relative payoffs of the players in the limit as $T\to 0$. From Eqn. (4.15) we find that

$$\begin{aligned}\lim_{T\to 0}\ \rho_1 &= \lim_{T\to 0}\left(\frac{e^{-rT}+(1-e^{-rT})e^{-\lambda T}-e^{-rT}m_1+e^{-(r+\lambda)T}(m_1-m_2)}{1-(1-e^{-rT})e^{-\lambda T}-e^{-rT}m_2-e^{-(r+\lambda)T}(m_1-m_2)}\right)\\ &= \frac{1-m_2}{1-m_1}\\ &= \mu_1/\mu_2.\end{aligned}$$

Now we have established the following theorem.

Theorem 4.2. *In the limit as $T\to 0$, the expected value of agreement in an SSPE converges to $\lambda/(\lambda+r)$, while the ratio of Player 1's and Player 2's SSPE utilities converges to μ_1/μ_2. The SSPE proposal of Player $i=1,2$ converges to $\theta_i=\mu_j\left(\frac{\lambda}{\lambda+r}\right)$. The expected implementation time in an SSPE converges to $1/\lambda$.*

The literature has established some findings on the SSPE of the canonical bargaining model that hold for a wide class of proposer selection protocols: In particular, agreement is reached immediately in every subgame, and in the limit as $\delta\to 1$, the SSPE proposals of all players converge to a common limit.[5]

In our model, there is a strictly positive expected delay on the equilibrium path. Hence, it is possible for the initial proposer's advantage to vanish although the proposals themselves do not converge to the fair split. This gap between our findings and the canonical bargaining model narrows if also the friction in communication is reduced, that is, if λ is large.

Corollary 4.2. *Suppose λ is sufficiently large. In the limit as $T\to 0$, the expected value of agreement in an SSPE is close to one, while the ratio of Player 1's and Player 2's*

[5]Different versions of these results appear, among others, in Banks and Duggan (2000), Kultti and Vartiainen (2010), Laruelle and Valenciano (2008), and Britz et al. (2010).

SSPE utilities is close to μ_1/μ_2. The SSPE proposal of Player $i = 1, 2$ to Player $j \neq i$ is close to μ_j. The expected implementation time is close to zero.

The corollary seems intuitive: As T becomes sufficiently small, and λ sufficiently large, the frictions which distinguish our bargaining game from the canonical bargaining model become negligible, and so our results collapse into the ones familiar from the canonical bargaining model.

This intuition, however, only holds true when considering the double limit "$\lim_{\lambda\to\infty} \lim_{T\to 0}$" of equilibrium variables. By contrast, Proposition 4.5 has shown that the surplus is split fairly with an expected implementation time of $2T$ in the limit as $\lambda \to \infty$ for any given T. Hence, at the double limit "$\lim_{T\to 0} \lim_{\lambda\to\infty}$" equilibrium is nearly efficient but also perfectly fair.

The interpretation is as follows: Consider the case where communication is noisy to a substantial degree, while implementation of agreements is possible almost at any time. Then, the equilibrium surplus allocation is determined by the distribution of proposal power in the same way as in the canonical bargaining model. Contrary to that model, however, there is a substantial expected delay of $1/\lambda$ on the equilibrium path of play, and the equilibrium proposals of the two players do not converge to a common limit.

Now consider the case where agreements can only be implemented at few and distant points in time, while noise in the communication is negligible. In that case, the surplus is split fairly, regardless of the distribution of proposal power. The expected equilibrium delay is $2T$.

4.5.2 Players' preferences over the length of bargaining rounds

The main trade–off can be summarized as follows: For $T > \ln(2)/\lambda$, every proposer offers the fair split, but the bargaining outcome is inefficient. A slight decrease in T improves the payoffs of both players. From the point $T = \ln(2)/\lambda$ onwards, any further gain in equilibrium efficiency will come at a cost in terms of fairness. Recall that for any $T \geq \ln(2)/\lambda$, the expected value of agreement is given by $\delta/(2-\delta)$, or equivalently, by $(2e^{rT} - 1)^{-1}$. At the point $T = \ln(2)/\lambda$, the expected value of agreement is $(2^b - 1)^{-1}$, where we denote $b = 1 + r/\lambda$. The quantity $(2^b - 1)^{-1}$ is the greatest expected value of agreement that can be reconciled with a perfectly fair split.

In the limit as $T \to 0$, the surplus is split according to the distribution of proposal power. However, the surplus is maximized in the limit as $T \to 0$, so fairness has a price in terms of efficiency.

Suppose that the proposal power of Player 1 is greater than that of Player 2, that is, $\mu_1 > 1/2$. As T decreases from $\ln(2)/\lambda$ towards zero, we should expect to see two effects: First, the surplus increases. Second, its distribution tilts more and more in favor of Player 1. Both effects are good for Player 1, so that Player 1 prefers T to be as small as possible.

For Player 2, however, the two effects work in opposite directions: Decreasing T hurts Player 2 because his share of the surplus becomes smaller, but on the other hand, Player 2 benefits because the surplus becomes bigger. Example 4.2 below illustrates these effects.

Example 4.2. *Suppose that $r = \lambda = 1$ so that $e^{-rT} = e^{-\lambda T} = e^{-T}$. Choosing $T \in (0, \ln(2)]$ amounts to choosing some $e^{-T} \in [\frac{1}{2}, 1)$. Whatever the proposer selection protocol may be, the expected value of agreement is*

$$e^{-T}v = e^{-T}/(1 + e^{-T}).$$

For the sake of this example, let us assume that $m_1 = 0.7$ and $m_2 = 0.3$. By substitution into Eqn. (4.13), we find that Player 1's SSPE payoff is

$$e^{-T}x_1 = \left(\frac{e^{-T}}{1 + e^{-T}}\right)\left(1.3e^{-T} - 0.6e^{-2T}\right).$$

This payoff is monotone increasing on the entire interval $e^{-T} \in [\frac{1}{2}, 1)$. Indeed, Player 1 becomes better off the smaller T is. In an analogous way, by substitution into Eqn. (4.14), we find that Player 2's SSPE payoff is

$$e^{-T}y_2 = \left(\frac{e^{-T}}{1 + e^{-T}}\right)\left(1 - 1.3e^{-T} + 0.6e^{-2T}\right).$$

This payoff is not monotonic in e^{-T} on the interval $[\frac{1}{2}, T)$. it attains a local minimum at $e^{-T} \approx 0.944$, where it evaluates to $e^{-T}y_2 \approx 0.149$. On the interval $[\frac{1}{2}, T)$, the optimal choice of e^{-T} for Player 2 is $e^{-T} = 1/2$, at which point Player 2's payoff evaluates to $e^{-T}y_2 = 1/6$.

The relation between the choice of T and the SSPE payoffs is more complex than the two effects discussed in the context of Example 4.2, however. In addition, we also need to take into account that the identity of the initial proposer has a non–monotonic effect on SSPE payoffs. The reason is as follows: We have assumed that Player 1 is the initial proposer in the first bargaining round. If $e^{-T} = 1/2$, all proposals are fair, so the gain from immediate agreement is split equally in each round. Hence, the identity of the initial proposer is irrelevant for the SSPE payoffs. As e^{-T} grows, however, there is a proposer advantage: The proposer in each round receives more of the gain from immediate agreement than the responder does. Due to time discounting, this effect has more bearing on the SSPE payoffs in earlier rounds than in later ones. Thus, there is an advantage to being the initial proposer. However, in the limit as e^{-T}, discounting becomes negligible, and so the premium for the initial proposer vanishes again. This is illustrated by Example 4.3 below.

Example 4.3. *Assume that $r = \lambda = 1$ and, moreover, $m_1 = m_2 = 0.9$. Also in this example, it is true that $e^{-rT} = e^{-\lambda T} = e^{-T}$, and that the expected value of agreement is $e^{-T}v = e^{-T}/(1+e^{-T})$. By substitution into Eqns. (4.13) and (4.14), we find that Player 1's SSPE payoff is*

$$e^{-T}x_1 = \left(\frac{e^{-T}}{1+e^{-T}}\right)\left(\frac{1.1e^{-T} - e^{-2T}}{1-0.8e^{-T}}\right),$$

and Player 2's SSPE payoff is

$$e^{-T}y_2 = \left(\frac{e^{-T}}{1+e^{-T}}\right)\left(\frac{1-1.9e^{-T} + e^{-2T}}{1-0.8e^{-T}}\right).$$

Player 1's SSPE payoff has a local maximum at $e^{-T} \approx 0.873$, where it evaluates to $e^{-T}x_1 \approx 0.306$. Player 2's SSPE payoff has a local minimum at $e^{-T} \approx 0.766$, where it evaluates to $e^{-T}y_2 \approx 0.147$. Player 2 is best off in the limit as $e^{-T} \to 1$, where he receives a payoff of $1/4$. At the point where $e^{-T} = 1/2$, Player 2 would only receive $1/6$.

In the canonical bargaining model, equilibrium offers become more fair when they can be made more frequently. In our model, the opposite may be true for some range of T: The distribution of proposal power has more and more bearing on the equilibrium allocation as $T \leq \ln(2)/\lambda$ becomes smaller.

In Example 4.3, the effect of the initial proposer's identity on SSPE payoffs is quite pronounced because the right to propose transitions from one player to the other only with a low probability of 0.1. Differently put, the initial proposer is likely to remain in that role for several rounds.

On the other extreme, it is also possible that the relation between T and the SSPE payoffs is determined only by the fact that smaller T implies more efficiency because other effects are quantitatively to weak to weigh in. This is illustrated by Example 4.4 below.

Example 4.4. *Assume that $r = \lambda = 1$ and, moreover, $m_1 = m_2 = 0$. Also in this example, it is true that $e^{-rT} = e^{-\lambda T} = e^{-T}$, and that the expected value of agreement is $e^{-T}v = e^{-T}/(1+e^{-T})$. By substitution into Eqns. (4.13) and (4.14), we find that Player 1's SSPE payoff is*

$$e^{-T}x_1 = \left(\frac{e^{-T}}{1+e^{-T}}\right)\left(\frac{2e^{-T} - e^{-2T}}{1+e^{-T}}\right),$$

and Player 2's SSPE payoff is

$$e^{-T}y_2 = \left(\frac{e^{-T}}{1+e^{-T}}\right)\left(\frac{1-e^{-T} + e^{-2T}}{1+e^{-T}}\right).$$

On the interval $e^{-T} \in [\frac{1}{2}, 1)$, SSPE payoffs of both players are monotone increasing. In addition, Player 1's SSPE payoff is concave in e^{-T} on that interval, while Player 2's SSPE payoff is convex on that interval. Both players obtain $1/6$ if $e^{-T} = 1/2$, and both obtain $1/4$ in the limit as $e^{-T} \to 1$.

The two previous examples have in common that the equilibrium split of the surplus is fair both at $T = \ln(2)/\lambda$ and in the limit as $T \to 0$. In those examples, it seems that there is no more trade–off between efficiency and fairness. So far, we have considered fairness only with regard to the ex ante *expected* split of the surplus. As discussed before, when T is small, SSPE proposals tend to be lopsided. So, the actual realized split of the surplus for small T is not fair "ex post." For instance, reconsider Example 4.4. In a round in which Player 1 is the proposer, his offer to Player 2 converges to $1/4$ in the limit as T goes to zero. It can be checked that, when Player 2 is the proposer, he offers only $1/4$ to Player 1. So the ultimate agreement will always allocate one player three times as much surplus as the other player.

4.5.3 Expected payoff to the final responder

In the previous subsection, we have considered which length of bargaining rounds is optimal for either of the two players, for various configurations of model parameters. Now we turn to the question how one might want to "design" an institutional environment in a way that takes both efficiency and fairness considerations into account. One simple way of doing this is to maximize the *expected payoff to the final responder.* This criterion is clearly responsive to both efficiency and fairness. Moreover, it takes into account not only the expected payoffs to the two players, but also the degree to which equilibrium proposals are lopsided.

We have shown that, in the limit as $T \to 0$, Player j is always offered $\mu_j \left(\frac{\lambda}{\lambda+r}\right)$ when he is the responder. If T is small enough, the probability that the offer which is eventually accepted is Player i's offer converges to μ_i. By symmetry, this means that the expected payoff to the final responder is $2\mu_1\mu_2 \left(\frac{\lambda}{\lambda+r}\right)$ in the limit as T goes to zero. Since $\mu_1+\mu_2 = 1$, we can also write it as $2(\mu_1 - \mu_1^2)\left(\frac{\lambda}{\lambda+r}\right)$.

If $T = \ln(2)/\lambda$, every proposer offers the fair split in an SSPE, and the expected value of agreement is $\frac{1}{2e^{rT}-1} = \frac{1}{2^b-1}$, where we recall that $b = 1 + r/\lambda$. Hence, the expected payoff to the final proposer is $\frac{1/2}{2^b-1}$ if $T = \ln(2)/\lambda$.

We see that the expected payoff to the final proposer is greater at $T = \ln(2)/\lambda$ than it is for sufficiently small T if and only if the following condition holds:

$$\frac{b}{2^b - 1} \geq 4(\mu_1 - \mu_1^2). \tag{4.16}$$

The right–hand side equals one if $\mu_1 = 1/2$, and is strictly less than one for any other $\mu_1 \in (0, 1)$. The left–hand side is monotone decreasing on the relevant interval $[1, \infty)$, it equals one if $b = 1$, and converges to zero in the limit as b goes to infinity.

If the proposal power of both players is equal, that is, $\mu_1 = \mu_2 = 1/2$, then the expected payoff to the final responder is greater in the limit as $T \to 0$ than it is if $T = \ln(2)/\lambda$. Note that this is true independently of the extent to which the equilibrium proposals are lopsided.

By contrast, if players differ ever so slightly in their proposal power, one can find parameter values r and λ so that the expected payoff of the final responder is greater at $T = \ln(2)/\lambda$ than it is for sufficiently small T.

4.6 Conclusion

The canonical bargaining model condenses each round into a single point in time. Offers are made, accepted, and implemented instantaneously. In the present paper, we have challenged this assumption, and proposed a bilateral bargaining model including frictions that the canonical approach abstracts away from. When bargaining rounds are of non–trivial length, players may make counter–offers. When communication is noisy, bargaining may fail not only due to disagreement, but also due to unsuccessful communication. When agreements can only be implemented according to a rigid time schedule, players may want to delay offers, and inefficiencies arise.

We have established conditions under which well–known results from the canonical bargaining model can be recovered in the limit as communication becomes less noisy, and the time schedule for implementing agreements becomes more flexible. This result can be interpreted as providing the conditions under which the canonical bargaining model is robust to the presence of frictions in each bargaining round.

One important question in bargaining theory is how conflict and delay can arise, while the canonical bargaining model strongly predicts immediate agreement in all subgames. One particularly simple explanation for conflict and delay could be noisy communication. There are many different potential interpretations of what we capture by modeling noisy communication: One very literal interpretation could be that players are uncertain about the frequency with which their opponents read their e–mail. An alternative interpretation is that contracts are often very complex, and it takes an uncertain amount of time to process, understand, and evaluate the exact terms of an offer. While our model gives an explanation for equilibrium delay, its predictions about the surplus allocation are still compatible with those of the canonical approach in the limit as the time schedule for implementing agreements is flexible enough. We have demonstrated that very noisy communication cannot only lead to prolonged conflict and delay in equilibrium, but also gives players an incentive to adopt tougher bargaining positions, which leads to more lopsided proposals and less predictable surplus allocations.

Another possible interpretation of the results in this paper is that they provide a rationale for institutions that constrain the time at which agreements can be implemented: When implementation of an agreement is subject to a rigid time schedule, this may make bargaining less efficient but render the surplus allocation more fair. In that sense, a more rigid bargaining institution may be useful to protect the weaker party in the negotiation from exploitation.

Appendix A

Proof of Proposition 4.1.

Since Player i accepts a counter–offer if and only if it offers him at least $\delta\widetilde{x}_i$, the best counter–offer that Player j can make is $\eta = \delta\widetilde{x}_i$. It immediately follows that Player j accepts an offer $\theta_i \geq 1 - \delta\widetilde{x}_i$ at any time. This explains why $\widehat{t}_j(\theta_i) = 0$ if $\theta_i \geq 1 - \delta\widetilde{x}_i$. Henceforth, suppose that $\theta_i < 1 - \delta\widetilde{x}_i$. Suppose that Player j makes his best counter–offer $\eta = \delta\widetilde{x}_i$ at time $(k-1)T + t$. Then, the probability that Player i receives the counter–offer before time kT is given by $1 - e^{-\lambda(T-t)}$. Making the optimal counter–offer at time $(k-1)T + t$ gives Player j an expected payoff of $(1 - e^{-\lambda(T-t)})(1 - \delta\widetilde{x}_i) + e^{-\lambda(T-t)}\delta\widetilde{y}_j$. By a standard argument, it is optimal for Player j to accept θ_i at time $(k-1)T + t$ if and only if

$$\theta_i \geq (1 - e^{-\lambda(T-t)})(1 - \delta\widetilde{x}_i) + e^{-\lambda(T-t)}\delta\widetilde{y}_j.$$

Rearranging this inequality, we obtain

$$e^{-\lambda(T-t)} \geq \frac{1 - \delta\widetilde{x}_i - \theta_i}{1 - \delta\widetilde{x}_i - \delta\widetilde{y}_j},$$

Solving for t yields

$$t \geq T - \left(\frac{1}{\lambda}\right) \ln\left(\frac{1 - \delta\widetilde{x}_i - \delta\widetilde{y}_j}{1 - \delta\widetilde{x}_i - \theta_i}\right).$$

Since t is non–negative by definition, the desired expression for $\widehat{t}_j(\theta_i)$ follows.

□

Proof of Proposition 4.2.

Suppose that Player i offers $1 - \delta\widetilde{x}_i - \varepsilon$ immediately, that is, at time $(k-1)T$. In that case, if $\varepsilon > 0$ is small enough, his expected payoff is

$$(1 - e^{-\lambda T})(\delta\widetilde{x}_i + \varepsilon) + e^{-\lambda T}\delta\widetilde{x}_i > \delta\widetilde{x}_i.$$

Hence, there is $\varepsilon > 0$ sufficiently small so that Player i can get an expected payoff of

$$\delta\widetilde{x}_i + (1 - e^{-\lambda T})\varepsilon > \delta\widetilde{x}_i.$$

Any choice of (θ_i, τ_i) which leads to an expected payoff of $\delta\widetilde{x}_i$ or less cannot be optimal. But if Player i offers $1 - \delta\widetilde{x}_i$, this leads to the payoff of $\delta\widetilde{x}_i$ no matter if the proposal is accepted or not. If Player i offers any $\theta_i > 1 - \delta\widetilde{x}_i$. at time $(k-1)T + \tau < kT$, the resulting expected payoff would be

$$(1 - e^{-\lambda(T-\tau)})(1 - \theta_i) + e^{-\lambda(T-\tau)}\delta\widetilde{x}_i < \delta\widetilde{x}_i$$

if Player j accepts, or $\delta\widetilde{x}_i$ if Player j does not accept. Finally, if Player i makes an offer only at time kT, this leads to a payoff of $\delta\widetilde{x}_i$ as well. $\square$

Proof of Proposition 4.3.

We show first that $\tau_i = \widehat{t}_j(\theta_i)$.

Let us suppose by way of contradiction that Player i chooses a pair (θ_i, τ_i) such that $\tau_i > \widehat{t}_j(\theta_i)$. Then the offer θ_i will be accepted if and only if it arrives before the deadline kT. Thus Player i's expected payoff is

$$(1 - e^{-\lambda(T-\tau_i)})(1-\theta_i) + e^{-\lambda(T-\tau_i)}\delta\widetilde{x}_i.$$

Consider a unilateral deviation under which Player i makes the same offer θ_i already at time $\tau_i - \varepsilon$ for some $\varepsilon > 0$ sufficiently small. This deviation is profitable because, by Proposition 4.2, we have that $1 - \theta_i > \delta\widetilde{x}_i$.

Now suppose that Player i chooses (θ_i, τ_i) such that $\tau_i < \widehat{t}_j(\theta_i)$. There are three cases to distinguish: First, if the offer arrives between times τ_i and $\widehat{t}_j(\theta_i)$, then Player j makes a counter–offer. Irrespective of whether or not this counter–offer arrives before time kT, the resulting expected payoff for Player i is $\delta\widetilde{x}_i$. Second, if the offer arrives between times $\widehat{t}_j(\theta_i)$ and kT, it is accepted, and so Player i receives $1 - \theta_i$. Third, if the offer does not arrive until time kT, then bargaining round k fails and so the expected payoff for Player i is $\delta\widetilde{x}_i$.

Thus Player i's expected payoff can be written as

$$e^{-\lambda(\widehat{t}_j(\theta)-\tau_i)}(1 - e^{-\lambda(T-\widehat{t}(\theta_i))})(1-\theta_i) + (1 - (e^{-\lambda(\widehat{t}_j(\theta_i)-\tau_i)})(1 - e^{-\lambda(T-\widehat{t}_j(\theta_i))}))\delta\widetilde{x}_i.$$

It is easy to verify that Player i has a profitable unilateral deviation by making the same offer θ_i at time $\widehat{t}_j(\theta_i)$ rather than at time τ_i, given that $1-\theta_i > \delta\widetilde{x}_i$. We can now conclude that an optimal choice of (θ_i, τ_i) by the proposer must be such that $\tau_i = \widehat{t}_j(\theta_i)$, as desired.

Now we show that in an SSPE, Player i chooses (θ_i, τ_i) such that $\theta_i = (1-e^{-\lambda(T-\tau_i)})(1-\delta\widetilde{x}_i) + e^{-\lambda(T-\tau_i)}\delta\widetilde{y}_j$. The argumemt is standard: Suppose first that $\theta_i > (1-e^{-\lambda(T-\tau_i)})(1-\delta\widetilde{x}_i) + e^{-\lambda(T-\tau_i)}\delta\widetilde{y}_j$. Then Player i has a profitable deviation in proposing $\theta_i - \varepsilon$ instead of θ_i for some $\varepsilon > 0$ sufficiently small. Now suppose that Player i offers some $\theta_i < (1-e^{-\lambda(T-\tau_i)})(1-\delta\widetilde{x}_i) + e^{-\lambda(T-\tau_i)}\delta\widetilde{y}_j$. Then Player j makes a counter–offer in which case Player i only gets $\delta\widetilde{x}_i$. But, by Proposition 4.2, we have that $1-\theta_i > \delta\widetilde{x}_i$.

□

Appendix B

In this Appendix, we provide the Mathematica code with which the solution to the system of Eqns. (4.5)–(4.12) can be verified. This code defines a total of eleven equations. The first eight of those correspond exactly to the aforementioned system of equations. The ninth equation in the code corresponds to the expected value of agreement. The tenth and eleventh equations are used to compute the equilibrium offers θ_1 and θ_2.

```
variables = {x1, x2, y1, y2, tx1, tx2, ty1, ty2, V, T1, T2};
equations = {x1 == d * tx1 + (1 - d * V) (p - p^2),
    x2 == d * tx2 + (1 - d * V) (p - p^2),
    y1 == V - x2,
    y2 == V - x1,
    tx1 == m1 * x1 + (1 - m1) y1,
    tx2 == m2 * x2 + (1 - m2) y2,
    ty1 == m2 * y1 + (1 - m2) x1,
    ty2 == m1 * y2 + (1 - m1) x2,
    V == (1 - p) / (1 - d * p),
    T1 == (1 - p) (1 - d * tx1) + p * d * ty2,
    T2 == (1 - p) (1 - d * tx2) + p * d * ty1};
eqsmatrixform = Normal[CoefficientArrays[equations, variables]];
```

normal | Arrays der Koeffizienten

```
M = eqsmatrixform[[2]];
v = -eqsmatrixform[[1]];
sol = LinearSolve[M, v];
```

löse lineare Gleichung

```
Print[MatrixForm[M], " ", MatrixForm[variables], " = ", MatrixForm[v]]
```

gib aus | Matritzenform | Matritzenform | Matritzenform

```
Print[MatrixForm[variables], " = ", MatrixForm[Simplify[sol]]]
```

gib aus | Matritzenform | Matritzenform | vereinfache

$$\begin{pmatrix}
1 & 0 & 0 & 0 & -d & 0 & 0 & 0 & d\left(p-p^2\right) & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & -d & 0 & 0 & d\left(p-p^2\right) & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\
-m1 & 0 & -1+m1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -m2 & 0 & -1+m2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
-1+m2 & 0 & -m2 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & -1+m1 & 0 & -m1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & d\ (1-p) & 0 & 0 & -d\,p & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & d\ (1-p) & -d\,p & 0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix} x1 \\ x2 \\ y1 \\ y2 \\ tx1 \\ tx2 \\ ty1 \\ ty2 \\ V \\ T1 \\ T2 \end{pmatrix}
=
\begin{pmatrix} p-p^2 \\ p-p^2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \frac{1-p}{1-d\,p} \\ 1-p \\ 1-p \end{pmatrix}$$

$$\begin{pmatrix} x1 \\ x2 \\ y1 \\ y2 \\ tx1 \\ tx2 \\ ty1 \\ ty2 \\ V \\ T1 \\ T2 \end{pmatrix}
=
\begin{pmatrix}
-\frac{(-1+p)\ (d+d\,m1\ (-1+p)+p-d\ (1+m2)\ p)}{(-1+d\ (-1+m1+m2))\ (-1+d\,p)} \\
-\frac{(-1+p)\ (d+d\,m2\ (-1+p)+p-d\ (1+m1)\ p)}{(-1+d\ (-1+m1+m2))\ (-1+d\,p)} \\
-\frac{(-1+p)\ (1-p+d\ (m1\ (-1+p)+p-m2\,p))}{(-1+d\ (-1+m1+m2))\ (-1+d\,p)} \\
\frac{(-1+p)\ (-1+p+d\ (m2+(-1+m1)\ p-m2\,p))}{(-1+d\ (-1+m1+m2))\ (-1+d\,p)} \\
\frac{(-1+p)\ (-1+m1+p-d\,p+(-2+d)\ m1\,p+d\,m2\,p)}{(-1+d\ (-1+m1+m2))\ (-1+d\,p)} \\
\frac{(-1+p)\ (-1+m2+p-d\,p+d\,m1\,p+(-2+d)\ m2\,p)}{(-1+d\ (-1+m1+m2))\ (-1+d\,p)} \\
-\frac{(-1+p)\ (m2+d\ (-1+m1+m2)\ (-1+p)+p-2\,m2\,p)}{(-1+d\ (-1+m1+m2))\ (-1+d\,p)} \\
-\frac{(-1+p)\ (m1+d\ (-1+m1+m2)\ (-1+p)+p-2\,m1\,p)}{(-1+d\ (-1+m1+m2))\ (-1+d\,p)} \\
\frac{-1+p}{-1+d\,p} \\
-\frac{(-1+p)\ \left(1+d^2\ (-1+m1+m2)\ p+d\ (-m2+p-2\,m1\,p)\right)}{(-1+d\ (-1+m1+m2))\ (-1+d\,p)} \\
-\frac{(-1+p)\ \left(1+d^2\ (-1+m1+m2)\ p+d\ (-m1+p-2\,m2\,p)\right)}{(-1+d\ (-1+m1+m2))\ (-1+d\,p)}
\end{pmatrix}$$

References

Banks, J.S. and J. Duggan (2000), A Bargaining Model of Collective Choice, *American Political Science Review, 94,* 73-88.

Britz, V., P.J.J. Herings and A. Predtetchinski (2010), Non–cooperative Support for the Asymmetric Nash Bargaining Solution, *Journal of Economic Theory, 145,* 1951–1967.

Fershtman, C. and D. Seidmann (1993), Deadline Effects and Inefficient Delay in Bargaining with Endogenous Commitment, *Journal of Economic Theory, 60,* 306–321.

Hart, S. and A. Mas-Colell (1996), Bargaining and Value, *Econometrica, 64,* 357–380.

Kultti, K. and H. Vartiainen (2010), Multilateral Non–cooperative Bargaining in a General Utility Space, *International Journal of Game Theory, 39,* 677–689.

Laruelle, A. and F. Valenciano (2008), Non–cooperative Foundations of Bargaining Power in Committees and the Shapley–Shubik Index, *Games and Economic Behavior, 63,* 341–353.

Kultti, K. and H. Vartiainen (2010), Multilateral Non–cooperative Bargaining in a General Utility Space, *International Journal of Game Theory, 39,* 677–689.

Laruelle, A. and F. Valenciano (2008), Noncooperative Foundations of Bargaining Power in Committees and the Shapley–Shubik Index, *Games and Economic Behavior, 63,* 341–353.

Ma, C. and M. Manove (1993), Bargaining with Deadlines and Imperfect Player Control, *Econometrica, 61,* 1313–1339.

Miyakawa, T. (2008), Non–cooperative Foundation of n–person Asymmetric Nash Bargaining Solution, *Journal of Economics of Kwansei Gakuin University, 62,* 1–18.

Ponsati, C. (1995), The Deadline Effect: A Theoretical Note, *Economics Letters, 48,* 281–285.

Rubinstein, A. (1982), Perfect Equilibrium in a Bargaining Model, *Econometrica, 50,* 97–110.

Chapter 5

Open Rule Legislative Bargaining

Joint with Hans Gersbach
Working Paper

Abstract: We consider non–cooperative bargaining on the division of a surplus under a simple majority rule. Bargaining takes place according to an "open rule" as originally suggested by Baron and Ferejohn (1989): Under an open rule, proposals can be amended before they are voted on. We first point out some gaps in Baron and Ferejohn's work, and provide a fresh analysis of open rule bargaining. We carefully distinguish between players endorsing a proposal and those voting in favor of a proposal. We devise a method to construct equilibrium candidates and to test whether these candidates are indeed equilibria. When players are sufficiently patient, we explicitly compute equilibrium outcomes. Compared to the canonical closed rule bargaining game, the equilibrium outcomes of open rule bargaining involve delays, but lead to more egalitarian surplus allocations. However, our results suggest that equilibrium delays tend to be longer, and surplus allocations tend to be less egalitarian than predicted by Baron and Ferejohn.

JEL classification: C72, C78, D72
Keywords: Bargaining, Legislatures, Open Rules, Baron and Ferejohn, Stationary Equilibrium

Acknowledgements: The authors would like to thank Tilman Börgers, Georgy Egorov, Juergen Eichberger, Hans Haller, Volker Hahn, Matthew Jackson, Roger Myerson, Bernhard Pachl, and seminar participants at ETH Zürich, Heidelberg University, and Princeton University for helpful comments. Tettje Halbertsma has provided excellent research assistance.

5.1 Introduction

We consider open rule bargaining on the division of a surplus under a simple majority rule. There is a vast game-theoretic literature on the resolution of surplus division problems through bargaining, see for instance the seminal papers by Rubinstein (1982) as well as Baron and Ferejohn (1989). The bargaining approach to surplus division problems has found important applications in political economy: The classical example is a parliament that negotiates about the allocation of funds from the government budget, while each member of the legislature wishes to obtain funds for projects in their own district. Such *legislative bargaining* models have been studied, among others, by Banks and Duggan (2000), Eraslan (2002), Kalandrakis (2004), Battaglini and Coate (2007), and Battaglini et al. (2014).

The game-theoretic bargaining literature emphasizes the importance of formal rules that structure the bargaining process. In particular, it assumes that negotiations consist of several rounds, where the rejection of a proposal triggers the next round. In these models, equilibrium outcomes depend crucially on who has the right to make a proposal at what point in time, and on how costly is the delay in moving from one round to the next.

Moreover, it matters how each round is structured. One important distinction can be made between "open rule" and "closed rule" bargaining. Under an *open rule,* a proposal may be amended one or several times before it is put to a vote. In contrast, under a *closed rule,* each proposal is immediately voted upon, allowing only Yes-or-No approval. A new proposal is only made when the previous one has been rejected. The seminal paper by Baron and Ferejohn (1989) offers a comparison of open and closed rule bargaining.

In practice, open rules are very common in legislative decision-making. For instance, open rule procedures are an important part of legislative bargaining in the U.S. Congress (Oleszek (2011)). In particular, the House of Representatives makes use of a variety of rules, including open rules, structured rules, and closed rules.[1] Open rules are not only prevalent at the federal level in the United States, but also at the state level (Primo (2003)).

However, most of the bargaining-theoretic literature has focused primarily on closed rules. In particular, this is true for most of the contributions following Baron and Ferejohn's seminal paper. The theoretical understanding of closed rule bargaining has been furthered considerably by the work of Ansolabehere et al. (2005), Banks and Duggan (2000, 2006), Diermeier and Feddersen (1998), Eraslan (2002), LeBlanc et al. (2000), and McCarty (2000), among others.

Compared to bargaining games with a closed rule, Baron and Ferejohn's model of

[1]Detailed information on the legislative processes in the House of Representatives can be found on the website of its Committee on Rules: https://rules.house.gov (retrieved on July 12th, 2019)

open rule legislative bargaining has received less attention. Primo (2007) shows how a proposer can randomize between different coalitions in open rule legislative bargaining. His findings indicate that the equilibrium found by Baron and Ferejohn cannot be unique. Fréchette et al. (2003) provide a theoretical and experimental investigation of open rules. Falconieri (2004) gives a comparative analysis of open and closed rules in the context of lobbying and delegation.

Arguably, one reason for the lack of literature on open rule bargaining is that the model proposed by Baron and Ferejohn (1989) is not easily tractable. Neither is it obvious what the right equilibrium concept for this model would be, nor is there a sharp equilibrium prediction as in the closed rule bargaining game. In the present paper, we address this gap in the literature: We propose a new approach to Baron and Ferejohn's model of open rule bargaining game. We introduce a suitable equilibrium refinement, which makes the analysis more tractable, and we obtain a sharp equilibrium prediction in the limit as the discount factor goes to one.

More specifically, we make the following contributions:

- We provide a rigorous definition of stationary strategies in the open rule legislative bargaining model. We define a class of equilibrium candidates that consists of relatively simple, and thus tractable, stationary strategies. We derive necessary and sufficient conditions under which such an equilibrium candidate is indeed a stationary subgame–perfect Nash equilibrium of the game.

- For the limit case, as the discount factor converges to one, we compute an explicit equilibrium prediction of the proposals, the payoffs, and the expected length of delay before an agreement.

- We compare the equilibrium outcome of open rule bargaining to that of the canonical closed rule bargaining process. While closed rule bargaining leads to immediate agreement, open rule bargaining typically leads to delays on the equilibrium path of play. We find that these equilibrium delays can be much longer than predicted by Baron and Ferejohn. We also show that the inefficiency inherent to an open rule can be so large that *all* players would be ex ante better off with a closed rule.

Our work is complementary to a stream of literature that analyzes bargaining with an endogenous status quo, see Anesi (2010), Diermeier and Fong (2011, 2012), and Bowen and Zahran (2012). In that class of bargaining models, negotiations continue even after an agreement has been reached. In each round, the status quo is given by the most recent agreement. This is different from bargaining under an open rule, where any agreement ends the game.

The rest of the paper is organized as follows: Section 2 contains the formal description of the open rule legislative bargaining game. In Section 3, we provide a rigorous definition of stationary strategies which is suitable for the analysis of this game. We provide a more detailed account of the relation between our present paper and Baron and Ferejohn's work in Section 4. In Sections 5 and 6, we conduct an equilibrium analysis of a slightly simplified version of the open rule bargaining game. In Section 7, we consider the limit as players are sufficiently patient and explicitly compute the equilibrium predictions. Afterwards, we use some numerical examples to illustrate our findings in Section 8. In Section 9, we argue that the main results and conclusions we have obtained for the simplified open rule bargaining game carry over to the original game. We offer some concluding remarks in Section 10. Most of the proofs are relegated to Appendix A. In Appendix B, we provide the Mathematica code that we have used to generate numerical examples.

5.2 The open rule legislative bargaining game

We consider open rule bargaining with an odd number of players, $n \geq 5$. The set of players is denoted by N, and we will frequently use i or j to index its members. There is a surplus of unit size to be divided among the players. Thus, the space of possible agreements is $\Delta^n = \{\theta \in \mathbb{R}^n_+ | \sum_{i \in N} \theta_i \leq 1\}$.[2] The decision to implement a particular surplus division is taken by simple majority voting, that is, it requires the approval of at least $(n+1)/2$ players. The bargaining process is structured in rounds $t = 0, 1, \ldots$. The number of rounds is potentially infinite.

Baron and Ferejohn's open rule bargaining process involves a complex chain of events in which players can make proposals, suggest amendments, choose between a proposal and an amendment, and eventually vote on the implementation of a proposal. In order to make this open rule bargaining process clear, we divide the description into the following three steps:

Step 1: Proposal on the floor

Consider any bargaining round t. Two cases must be distinguished: Either, bargaining round t begins with a *proposal on the floor,* or it begins *without a proposal on the floor.*

- *Proposal on the floor:* If round t begins with a proposal on the floor, then round t of the game directly proceeds to Step 2 below.

- *No proposal on the floor:* If round t begins without a proposal on the floor, then a *proposer* is randomly chosen from N with equal probability. Let us say that player i is chosen as the proposer. Then, player i chooses some proposal $\overline{\theta} \in \Delta^n$, which thereby becomes the proposal on the floor.[3] Now the game proceeds to Step 2.

[2] Without loss of generality, we will focus on feasible agreements that satisfy $\sum_{i \in N} \theta_i = 1$. In particular, a proposer never finds it optimal to make a proposal that does not fully divide the available surplus.

[3] Here, we renounce indexing $\overline{\theta}$ by i to ease presentation.

Bargaining round t can only begin with a proposal on the floor if that proposal has been made in a previous round. Therefore, the initial bargaining round $t = 0$ begins without a proposal on the floor.

Step 2: Amendment or endorsement

Suppose that the proposal $\overline{\theta} \in \Delta^n$ made by some player $i \in N$ is on the floor in round t. Now, a new proposer is randomly chosen with equal probability from $N \setminus \{i\}$. Let us say that player j has been chosen. Then, player j decides whether to *endorse* or *amend* the proposal on the floor.

- *Endorsement:* If player j *endorses* the proposal on the floor $\overline{\theta}$, then round t of the game proceeds directly to Step 3.

- *Amendment:* Suppose that player j chooses to amend the proposal $\overline{\theta}$ on the floor. He does so by announcing an *amendment* $\theta' \in \Delta^n \setminus \{\bar{\theta}\}$. Then, all players simultaneously cast votes in favor of the proposal on the floor $\bar{\theta}$ or in favor of the amendment θ'. If at least $(n+1)/2$ players vote in favor of θ', then bargaining round $t+1$ begins with the amendment θ' as the new proposal on the floor. If at least $(n+1)/2$ players vote in favor of $\overline{\theta}$, then bargaining round $t+1$ begins with the proposal $\overline{\theta}$ on the floor. Note that players can keep making amendments, and thus repeating Step 2, indefinitely. However, a new bargaining round begins every time a new amendment is made.

Step 3: Voting on an endorsed proposal

Now suppose that in some bargaining round t, a proposal on the floor $\overline{\theta}$ is endorsed by player j. Then, all players simultaneously cast votes in favor or against the endorsed proposal $\overline{\theta}$. Again, there are two cases:

- *Majority approval:* If at least $(n + 1)/2$ players accept the endorsed proposal $\overline{\theta}$, then the game ends and $\overline{\theta}$ is implemented.

- *No majority approval:* If strictly less than $(n + 1)/2$ players accept $\overline{\theta}$, then the game moves to round $t + 1$. That bargaining round begins again in Step 1, without a proposal on the floor.

Note that a new bargaining round begins whenever either *(i)* an amendment is made (in Step 2), or *(ii)* an endorsed proposal is not approved by the majority (in Step 3). Every time a new bargaining round starts, a discount factor $\delta \in (0, 1)$ is applied. This discount factor can be suitably interpreted as a measure for the players' impatience. Furthermore, we assume that players are risk-neutral. Thus, if a proposal $\theta \in \Delta^n$ is implemented in bargaining round t, then player i receives a payoff of $\delta^t\theta_i$. If no proposal is ever endorsed, or if no endorsed proposal is ever approved by the majority, then bargaining is trapped in *perpetual disagreement,* which gives all players zero payoffs. This completes the formal description of the open rule legislative bargaining game (henceforth ORBG) $G(\delta, n)$. It corresponds to the open rule bargaining game originally proposed by Baron and Ferejohn (1989).

Throughout most of this paper, we are going to analyze a slightly abridged version of the ORBG, which we call the *simplified ORBG,* and denote by $\widehat{G}(\delta, n)$. In Section 9, we will argue that the main results and conclusions derived for the simplified ORBG also hold in the ORBG itself. The simplified ORBG differs from the ORBG as follows: Whenever a player makes an amendment to a proposal on the floor in Step 2, the amendment immediately replaces the proposal on the floor, without a vote being held.

We make two remarks for a better understanding of bargaining power in this game:

First, consider a history of this game where players are in Step 3 and thus vote on an endorsed proposal. Their choice is either to stop bargaining and implement the proposal now, or to move back to Step 1 and start bargaining from scratch in the next round. This is similar to the choice that players make when responding to proposals in a closed rule bargaining game. In such games, the prospect of discounting discourages players from rejecting a proposal, which leads to a bargaining advantage for the proposer, often called the *proposer premium.* With an open rule, it seems intuitive that this proposer premium is shared between the player who has originally made the proposal, and the one who has endorsed it. One crucial question in this paper will be how many players share the proposer premium, and how it is divided.

Second, consider a history of the game where players vote between a proposal on the floor and an amendment. They decide whether the current proposal on the floor remains the proposal on the floor in the next round, or whether the amendment becomes the new proposal on the floor. Regardless of the outcome of such a vote, another round of bargaining is required to reach an agreement, and thus another round of discounting will occur either way. Loosely speaking, while players are under "time pressure" when they vote on an endorsed proposal, this time pressure does not affect them when choosing between a proposal on the floor and an amendment.

5.3 Stationary strategies

It is well-known that non-cooperative bargaining games with more than two players admit a wide multiplicity of subgame-perfect equilibrium allocations. The bargaining literature has focused on analyzing subgame-perfect equilibria in stationary strategies (SSPE). In closed rule bargaining games in the tradition of Rubinstein (1982), the meaning of a "stationary" strategy is straightforward: Each player makes the same proposal every time he is the proposer. Moreover, each player uses the same acceptance rule whenever he is a responder. In this paper, however, we depart from most of the bargaining literature by considering an open rule. In the open rule bargaining game, the appropriate definition of a "stationary" strategy is less obvious. The purpose of the present section is to define the stationary strategies that we use in our analysis. These strategies are "mixed" in the following sense: They allow players to choose lotteries over different actions only when they are making proposals or amendments. The purpose of these lotteries is to ensure that the proposer can treat all other players equally. That is, a proposer can choose a configuration of payoffs that he wants to offer to the other players, while leaving it to chance which payoff is offered to which player.

In order to make this idea more precise, let us define an *anonymous proposal* as a vector $\eta \in \Delta^n$ such that $\eta_1 \geq \eta_2 \geq \ldots \geq \eta_n$. Furthermore, let $\mathcal{M}(i)$ be the collection of $(n \times n)$-permutation matrices M such that $m_{1,i} = 1$. For any anonymous proposal $\eta \in \Delta^n$, let

$$\Theta^i(\eta) = \{\theta \in \Delta^n |\ \theta = M^\top \eta \text{ for some } M \in \mathcal{M}(i)\}.$$

Each of the proposals in $\Theta^i(\eta)$ assigns η_1 to player i, and the payoffs $\eta_2, \ldots, \eta_n$ to the remaining players. Of course, the assignment of the payoffs to the individual players $N \setminus \{i\}$ differs across the different elements of $\Theta^i(\eta)$.[4]

Formally, in the open rule legislative bargaining game $G(\delta, n)$, a *stationary strategy* for player i consists of the following four elements:

1. An *anonymous proposal* $\eta^i \in \Delta^n$, such that at every history at which there is no proposal on the floor and player i is the proposer, he randomizes uniformly among all the elements of $\Theta^i(\eta^i)$.

2. Let $P(\Delta^n)$ denote the power set of Δ^n. An *amendment rule* is a map $\psi^i : \Delta^n \times N \setminus \{i\} \to P(\Delta^n)$ which prescribes how player i behaves when he is the proposer and proposal θ, made by player k, is on the floor. If $\psi^i(\theta, k) = \{\theta\}$, player i endorses player k's proposal. If $\psi^i(\theta, k) = \{\theta'\}$ for some $\theta' \neq \theta$, player i makes an amendment θ' when k's proposal θ is on the floor. If $\psi^i(\theta, k)$ consists of $m \geq 2$ elements and $\theta \in \psi^i(\theta, k)$, then player i endorses k's proposal θ with probability $1/m$, and chooses every element of $\psi^i(\theta, k) \setminus \{\theta\}$ as an amendment with probability $1/m$. If $\psi^i(\theta, k)$ consists of at least two elements, and $\theta \notin \psi^i(\theta, k)$, then player i amends player k's proposal θ. In particular, he chooses the amendment from $\psi^i(\theta, k)$ uniformly at random.

[4] The requirement that player i randomize uniformly among all the elements of $\Theta^i(\eta)$ adds an anonymity requirement to the stationarity requirement. Strategies in which proposals are made in a "stationary but not anonymous" way play no role in our paper. They would complicate the analysis without offering new insights. Therefore, it seems convenient to include the anonymity requirement in the definition of stationarity.

3. A *selection rule* $\chi^i : \Delta^n \times \Delta^n \times N \times N \to \{\text{Proposal, Amendment}\}$ indicates player i's behavior at histories where he votes between a proposal and an amendment. This voting decision can be conditioned on the proposal on the floor, on the amendment, and on the identities of the players who have made the proposal on the floor and the amendment.

4. An *acceptance rule* $A^i \subset \Delta^n$ describes player i's voting decisions at histories where he votes on an endorsed proposal. More precisely, player i votes in favor of an endorsed proposal θ if and only if $\theta \in A^i$. Of course, the set A^i is specified independently of the history of play.

We use the following notation to describe the *stationary strategy* for player i:

$$\sigma^i = (\eta^i, A^i, \psi^i, \chi^i),$$

and we write $\sigma = (\sigma^1, \ldots, \sigma^n)$ for a profile of stationary strategies. We note that our definition of stationary strategies implies the following: When there is no proposal on the floor, a proposer can only make an anonymous proposal. Amendments, however, need not be anonymous. The reason is that we want to allow an amendment to condition on the identity of the player who has made the proposal on the floor. More specifically, we will be interested in amendments which permute the amounts offered to the current and previous proposers, while leaving the remaining $n - 2$ components of the proposal unchanged.

Given this definition of a stationary strategy, the equilibrium concept is perfectly standard: Indeed, a *stationary subgame-perfect equilibrium (SSPE)* is a profile of stationary strategies that is a subgame-perfect equilibrium.

In the simplified ORBG, a stationary strategy consists of an anonymous proposal, an amendment rule, and an acceptance rule, while a selection rule is redundant.[5]

5.4 Relation to Baron and Ferejohn

Although the present paper deals with the open rule bargaining model proposed by Baron and Ferejohn (1989), our analysis and results differ from theirs in several respects. In this section, we discuss the fundamental reasons for these differences:

1. Baron and Ferejohn's equilibrium analysis lacks a comprehensive description of the relevant strategy profiles. In particular, their equilibrium strategies are not fully

[5] Mutatis mutandis, the definition of an SSPE in the simplified ORBG corresponds to that in the ORBG, to which we will return in Section 9.

specified off the path of play. This is problematic at histories where players vote on whether or not to replace the current proposal on the floor with an amendment. More specifically, Baron and Ferejohn impose that a player votes in favor of the amendment if he is "indifferent" between the proposal on the floor and the amendment. Unfortunately, it is not straightforward what it means to be indifferent between the proposal on the floor and the amendment: Player i's preferences over the proposal on the floor, say $\overline{\theta}$, and the amendment, say θ', do not only depend on the components $\overline{\theta}_i$ and θ'_i, but also on the probabilities with which either $\overline{\theta}$ or θ' will be endorsed or amended in the future. For instance, even if $\bar{\theta}_i < \theta'_i$ player i may want to vote in favor of $\bar{\theta}$ because he believes that $\bar{\theta}$ will be endorsed with a higher probability than θ'. Along a path of play of Baron and Ferejohn's supposed SSPE, a proposal on the floor and an amendment always have the same probability of being endorsed. However, this is no longer true off the equilibrium path. In this paper, we work around this problem in two different ways: First, we analyze the simplified ORBG in which the problem is redundant. Second, in Section 9, we return to the original ORBG and show that players' best-responses to deviations from SSPE must have a certain recursive structure. Therefore, we can do an equilibrium analysis without explicitly determining the optimal voting behavior for each player, for each proposal on the floor, and a for each amendment. This analysis confirms that the results and conclusions obtained in the simplified ORBG carry over to the ORBG itself.

2. Baron and Ferejohn tacitly assume that a player who is willing to vote for a given proposal is also willing to endorse it. However, we will demonstrate that this need not be true: We will see that there may be players who would want to amend a proposal if they had the chance to do so, but who would nevertheless want to vote in favor of that same proposal once it had been endorsed.[6] Taking this possibility into account changes some of the analysis and conclusions. In particular, we find longer equilibrium delays and less egalitarian allocations than Baron and Ferejohn.

3. In the present paper, we look at a class of stationary equilibria with the following property: On the equilibrium path, whenever a player amends a proposal, he does so while leaving unchanged the payoffs offered to all players other than himself and the player who has made the previous proposal. Put another way, an amendment

[6] This problem with Baron and Ferejohn's analysis was recognized earlier in a working paper by Fahrenberger and Gersbach (2007). In the present paper, we adopt a new approach that differs from the one in Baron and Ferejohn (1989) as well as from Fahrenberger and Gersbach (2007): We focus on a class of stationary equilibria that involves particularly simple amendment rules which we will call "simple swap." This makes the problem more tractable than in any previous work we are aware of. Based on our new approach, we can construct and test equilibrium candidates for any values of the model parameters. Within the class of SSPE we consider, we can explicitly compute the limit of equilibrium payoffs as $\delta \to 1$ and the equilibrium number of players who endorse a proposal.

merely permutes two components of a proposal. We will call this kind of amendment a *simple swap.* Our focus on this class of stationary equilibria differs from the approach taken by Baron and Ferejohn. It allows us to express the proposals and payoffs associated with the equilibrium candidates as solutions to a relatively simple and tractable system of linear equations.

The relation of the present paper to Baron and Ferejohn's work can be summarized as follows: We point out that the strategy profiles which they claim to be equilibria are not fully specified. We have found it impossible to write down equilibrium strategies which generally support Baron and Ferejohn's supposed equilibrium payoffs. Whether such equilibrium strategies exist at all remains an open question. Our analysis focuses on an alternative class of stationary equilibria that is more easily tractable. Based on this class of stationary equilibria, we obtain results that qualify some of Baron and Ferejohn's conclusions.

5.5 Equilibrium candidates for the simplified ORBG

In this section, we entirely focus on the simplified ORBG, and discuss a particular family of stationary strategy profiles that we call *k-candidates with simple swaps.* Such a stationary strategy profile has the following properties:

- On the path of play induced by a k-candidate with simple swaps, whenever a player amends a proposal on the floor, he does so by simply swapping his component of the proposal on the floor with that of the player who has made the proposal on the floor.

- Every proposal and every amendment made on a path of play of a k-candidate with simple swaps has the following structure: The proposer offers k players a payoff that makes them willing to endorse the proposal, and to vote in its favor. If $k \leq \frac{n-1}{2}$, the proposer offers an additional $\frac{n-1}{2} - k$ players a payoff that makes them willing to vote in favor of the proposal once it has been endorsed, but not to endorse it themselves.

Take the number of players n and the discount factor δ as given. For any $k = 1, \ldots, n-1$, let the quadruple (V_k, W_k, X_k, Y_k) be defined as the solution to the following system of equations:

$$
\begin{aligned}
V_k &= \left(\frac{k}{n-1}\right) X_k + \left(\frac{n-1-k}{n-1}\right) \delta W_k, && (5.1)\\
W_k &= \frac{Y_k - V_k}{n-1}, && (5.2)\\
X_k &= 1 - k\delta V_k - \max\left\{0, \frac{n-1}{2} - k\right\}\left(\frac{\delta}{n}\right) Y_k, && (5.3)\\
Y_k &= \frac{k}{\delta k + (1-\delta)(n-1)}. && (5.4)
\end{aligned}
$$

We will show below that a solution to this system exists. Before proceeding to the formal definition of the equilibrium candidates, let us intuitively describe how the variables (V_k, W_k, X_k, Y_k) enter the construction of a k-candidate with simple swaps: The proposer always offers the amount X_k to himself. If his proposal is endorsed, it will also be accepted, and so the proposer will receive X_k. If the proposal on the floor is not endorsed, then the current proposer will certainly not be the proposer again in the next round. In this case, he will receive the expected payoff of a responding player, which is W_k, after one round of discounting. Thus, the proposer's expected payoff, denoted by V_k, must be a weighted average of X_k and δW_k.

The weight given to X_k is the probability that the proposal is endorsed. This probability equals $k/(n-1)$ because the proposer gives k out of the $n-1$ responding players an incentive to endorse the proposal. It will turn out that δV_k is the amount that makes a responding player exactly indifferent between endorsing the proposal on the floor and making an amendment. Indeed, it is the expected payoff from making an amendment: A player who makes an amendment triggers one round of discounting, and then takes the place of the proposer. These considerations explain Eqn. (5.1).

It will turn out that a player who is willing to endorse a proposal is also willing to vote in its favor once it has been endorsed. Therefore, if $k \geq (n-1)/2$, a majority of players votes in favor of the endorsed proposal. If $k \leq (n-3)/2$, however, the proposer and the k players willing to endorse the proposal on the floor do not form a majority. Hence, the proposer must convince $\frac{n-1}{2} - k$ additional players to vote for the proposal on the floor once it has been endorsed. The expected continuation payoff for any player after the rejection of an endorsed proposal is $\frac{\delta}{n}(V_k + (n-1)W_k)$. Writing Y_k for the quantity $V_k + (n-1)W_k$ as in Eqn. (5.2), this explains Eqn. (5.3) above. All that remains to be explained is Eqn. (5.4). The quantity Y_k is the sum of the expected payoffs to all players. Equivalently, it can be thought of as the value of the surplus discounted by the expected delay. If the very first proposal which is made is endorsed, the total surplus of size one is divided. If t amendments are made before the t^{th} amendment is endorsed, the surplus divided is of size δ^t. Since each proposal is endorsed with probability $k/(n-1)$, the probability that the t^{th} amendment is endorsed is $\left(1-\frac{k}{n-1}\right)^t\left(\frac{k}{n-1}\right)$. We note that $1-k/(n-1)<1$ since $1\leq k\leq n-1$.

Summing over t from zero to infinity and rewriting yields the expression in Eqn. (5.4).

For the analysis in the remainder of the paper, an important auxiliary result is that the variables (V_k, W_k, X_k, Y_k) as defined in Eqns. (5.1)-(5.4) are strictly positive. The formal claim is stated in Proposition 5.1 below. The proof is provided in Appendix A.

Proposition 5.1. *The system of Eqns. (5.1)-(5.4) has a unique solution. Furthermore: (i) If $k=n-1$ and $\delta=1$, then $V_k=W_k>0$. (ii) For any other choices of $k=1,\dots,n-1$ and $\delta\in(0,1]$, it holds that $V_k>W_k>0$. (iii) For any $k=1,\dots,n-1$, all components of solutions (V_k, W_k, X_k, Y_k) to the system of equations (5.1)-(5.4) are strictly positive.*

To provide a formal definition of the k-candidate with simple swaps, we need the following notation: For any proposal $\theta\in\Delta^n$ and any two players $i,j\in N$, let $\pi_{i\leftrightarrow j}(\theta)$ be the permutation of θ which swaps components i and j, while leaving all other components unchanged. Let $H^{\emptyset}$ be the set of histories at which there is no proposal on the floor. Let H^{θ} be the set of histories at which the proposal θ is on the floor. In particular, let $H^{i,\theta}\subset H^{\theta}$ be the set of histories at which the proposal on the floor is θ, and the author of that proposal is player i. Finally, let $H^{i,\theta}_j\subset H^{i,\theta}$ be the set of histories at which the proposal on the floor is θ, the author of that proposal is player i, and the current proposer is player j.

We now provide a formal definition of the k-candidate with simple swaps:

Definition 5.1. *Consider the simplified ORBG $\widehat{G}(\delta,n)$. Let (V_k, W_k, X_k, Y_k) be defined as solutions to Eqns. (5.1)-(5.4). For every $k=1,\dots,n-1$, a profile of stationary strategies is a k-*candidate with simple swaps* if the following hold:*

1. *At every history $h\in H^{\emptyset}$, the proposer makes an anonymous proposal η which gives himself X_k, gives δV_k to k other players, gives $\frac{\delta}{n}((n-1)W_k+V_k)$ to $\max\{0,\frac{n-1}{2}-k\}$ more players, and zero to all remaining players.*

2. *Consider a history $h\in H^{i,\theta}_j$. At such a history, the proposer is $j\in N$, and the proposal on the floor θ was made by player $i\in N$. Suppose that $\theta\in\Theta^i(\eta)$, where η is the anonymous proposal described in Point 1. Player j endorses the proposal on the floor θ if and only if $\theta_j\geq\delta V_k$. Otherwise, he makes the amendment $\pi_{i\leftrightarrow j}(\theta)$. Now suppose that $\theta\notin\Theta^i(\eta)$. Player j endorses θ if and only if there are at least $(n-1)/2$ players l in $N\setminus\{i\}$ with $\theta_l\geq\frac{\delta}{n}((n-1)W_k+V_k)$ and, moreover, it holds that $\theta_j\geq\delta V_k$. Otherwise, player j randomly chooses an amendment from $\Theta^j(\eta)$ with equal probability.*

3. *Whenever player i votes on an endorsed proposal θ, he votes in favor if and only if $\theta_i \geq \frac{\delta}{n}((n-1)W_k + V_k)$.*

Note that the three points of Definition 5.1 above correspond to the elements of a stationary strategy as defined in Section 3: Point 1 describes the *anonymous proposal,* Point 2 specifies the *amendment rule,* and the *acceptance rule* is spelled out in Point 3. Recall that, since we are considering the simplified ORBG, it is redundant to specify a selection rule.

Since $V_k \geq W_k$, according to Proposition 5.1, we have $\frac{\delta}{n}Y_k \leq \delta V_k$, and thus a player who is willing to endorse a proposal will also vote in its favor. Finally, we conclude from Proposition 5.1 above that a proposer offers the highest share of surplus to himself. Indeed, we obtain the following corollary:

Corollary 5.1. *(i) If $k = n-1$, then $X_k = V_k$. (ii) For any other choices of $k = 1, \ldots, n-2$ and $\delta \leq 1$, it holds that $X_k > \delta V_k$.*

The proof of Corollary 5.1 can be found in Appendix A.

By restricting attention to the simplified ORBG and to k-candidates with simple swaps, the analysis of stationary strategy profiles becomes more tractable than in any previous work on open rule bargaining that we know of. There are two reasons for this: :

- In a simplified ORBG, there is a strategic equivalence between subgames that start at a node where a proposal on the floor can be amended or endorsed, and subgames that start at a node where no proposal is on the floor.

- If a k-candidate with simple swaps is played, the actions taken after a history $H_j^{i,\theta}$ do not depend on whether the proposal θ was originally made as an amendment to some other proposal, or whether it was made at a history without a proposal on the floor.

Intuitively, in a k-candidate with simple swaps played in a simplified ORBG a player who can make an amendment to a proposal on the floor can achieve the same payoff (up to discounting) that he could also achieve if he were the proposer at a history without a proposal on the floor.

5.6 Testing equilibrium candidates

In this section, we introduce a test to verify whether a k-candidate with simple swaps is an SSPE of a simplified ORBG.

Proposition 5.2. *A k-candidate with simple swaps is an SSPE of the simplified ORBG if and only if there is no profitable unilateral deviation from it at any history $h \in H^{\emptyset}$.*

Proof. It is easily verified that a profitable unilateral deviation from a k-candidate with simple swaps is impossible at histories where players vote on an endorsed proposal. Thus, we have to focus on the possibility of profitable unilateral deviations from k-candidates with simple swaps at histories where a proposal can be made. Recall that a proposal can be made at histories in $H^{\emptyset}$ or at histories in H^{θ} through an amendment. Consider a history $h \in H^{\theta}$ at which player i chooses to endorse or amend the proposal on the floor θ. Suppose that player i obtains an expected payoff of $\delta\widetilde{V}$ if he makes the amendment $\widetilde{\theta}$. Now consider a history in $H^{\emptyset}$ where player i is the proposer. At that history, player i can obtain a payoff of $\widetilde{V}$ by proposing $\widetilde{\theta}$. When player i proposes at a history in $H^{\emptyset}$, his expected payoff is V_k, and when he proposes at a history in H^{θ}, his expected payoff is δV_k. Thus, if player i has a profitable deviation at a history in H^{θ}, then he also has a profitable deviation at a history in $H^{\emptyset}$. $\square$

Proposition 5.2 shows that we can apply the one-shot deviation principle in order to test whether a k-candidate with simple swaps is an SSPE in the simplified ORBG. Consequently, in order to test whether a k-candidate with simple swaps is an SSPE in the simplified ORBG, we only have to consider profitable unilateral deviations at histories in $H^{\emptyset}$.

In the simplified ORBG, all histories at which a particular player can make a proposal or an amendment are "equivalent" in the sense that the continuation game is the same.

Next, we consider deviations from the k-candidate with simple swaps. Suppose that player i makes a unilateral one-shot deviation from the k-candidate with simple swaps by proposing the amount δV_k to $k+1$ instead of k players, proposing $\left(\frac{\delta}{n}\right) Y_k$ to $\max\left\{0, \frac{n-1}{2} - (k+1)\right\}$ players, and proposing to take the remainder for himself. We denote the proposer's expected gain from such a deviation by λ_k^+. Similarly, let λ_k^- denote the proposer's expected gain from a unilateral one-shot deviation under which the proposer offers the amount δV_k only to $k-1$ instead of to k players, and offers $\left(\frac{\delta}{n}\right) Y_k$ to $\max\left\{0, \frac{n-1}{2} - (k-1)\right\}$ players. In order to understand the expressions below, recall that Point 2 in Definition 5.1 says that if after the deviation, the proposal is amended, then that amendment is again based on the anonymous proposal associated with the k-candidate with simple swaps.

$$\lambda_k^+ = \begin{cases} 0 & \text{if } k = n-1, \\ -\left(\frac{k+1}{n-1}\right)\delta V_k + \left(\frac{1}{n-1}\right)(X_k - \delta W_k) & \text{if } k \in \{(n-1)/2, \ldots, n-2\}, \\ -\left(\frac{k+1}{n-1}\right)\left(\delta V_k - \delta\left(\frac{1}{n}\right)Y_k\right) + \left(\frac{1}{n-1}\right)(X_k - \delta W_k) & \text{if } k \in \{1, \ldots, (n-3)/2\}. \end{cases} \tag{5.5}$$

$$\lambda_k^- = \begin{cases} \left(\frac{k-1}{n-1}\right)\delta V_k - \left(\frac{1}{n-1}\right)(X_k - \delta W_k) & \text{if } k \in \{(n+1)/2, \dots, n-1\}, \\ \left(\frac{k-1}{n-1}\right)\left(\delta V_k - \delta\left(\frac{1}{n}\right)Y_k\right) - \left(\frac{1}{n-1}\right)(X_k - \delta W_k) & \text{if } k \in \{2, \dots, (n-1)/2\}, \\ 0 & \text{if } k = 1. \end{cases} \tag{5.6}$$

It is straightforward that the k-candidate with simple swaps can only be an SSPE if λ_k^+ and λ_k^- are non-positive. The next proposition implies the converse: If the proposer has any profitable deviation, then either λ_k^+ or λ_k^- must be strictly positive. In particular, if the proposer cannot gain by offering δV_k to one additional player, or to one player less, then he cannot gain either by offering δV_k to any number of players other than k.

Proposition 5.3. *If there exists $\widehat{\theta} \in \Delta^n$ such that proposing $\widehat{\theta}$ instead of the proposal prescribed by the k-candidate with simple swaps is a profitable deviation for the proposer, then either $\lambda_k^+ > 0$ or $\lambda_k^- > 0$.*

The proof of Proposition 5.3 is relegated to Appendix A. Propositions 5.2 and 5.3 lead to the following theorem:

Theorem 5.1. *A k-candidate with simple swaps is an SSPE of the simplified ORBG if and only if $\lambda_k^+ \leq 0$ and $\lambda_k^- \leq 0$.*

Baron and Ferejohn (1989) have already argued that a proposer should choose a large k when δ is small. The intuition is as follows: For small δ, any delay is very costly. Hence, it seems intuitive that the proposer finds it optimal to ensure immediate endorsement and acceptance of his proposal. In order to ensure that his proposal is endorsed immediately with probability one, he needs to make all other players willing to endorse it. The proposition below formalizes this argument. The proof is given in Appendix A.

Proposition 5.4. *If δ is sufficiently small, then the $(n-1)$-candidate with simple swaps is an SSPE.*

In an $(n-1)$-candidate with simple swaps, immediate agreement is reached on an allocation which gives $\frac{1}{1+\delta(n-1)}$ to the proposer and $\frac{\delta}{1+\delta(n-1)}$ to each of the other players. This corresponds exactly to the payoff division that one would expect under closed rule unanimity bargaining.[7]

It is important to emphasize that our analysis so far does not yield results on the "existence" or "uniqueness" of k-candidates that are SSPE in the simplified ORBG. Without any restrictions on the parameters δ and n, we do not claim that there must be a k so

[7] A closed rule bargaining game with linear utility functions and equal recognition probabilities is a special case of the games studied in Laruelle and Valenciano (2008) and Britz et al. (2014).

that the k-candidate with simple swaps is an SSPE. We do not show either that there is at most one k so that the k-candidate with simple swaps is an SSPE. In Section 8, however, we do consider some numerical examples. In each of the examples, it does turn out that exactly one k-candidate with simple swaps is an SSPE.

In the next section, we consider k–candidates with simple swaps that are SSPE in the limit as $\delta \to 1$, and provided $n \geq 15$. In that case, we do obtain results which show that there is a unique k such that the k–candidate with simple swap is an SSPE.

5.7 Stationary equilibrium with patient players

So far, we have defined a family of equilibrium candidates in the simplified ORBG, and we have introduced a test to verify which of these candidates are indeed SSPE of the simplfied ORBG. In the present section, we will focus on the case where the discount factor is sufficiently close to one. In that case, we will explicitly compute the limit of SSPE payoffs.

As a first step, we show that for sufficiently large δ and n, a k-candidate with simple swaps can only be an SSPE if $k \leq (n-3)/2$. Consequently, for sufficiently large δ and n, a k-candidate with simple swaps can only be an SSPE if there are players who are willing to vote in favor of proposals that they would not be willing to endorse.

We say that a k-candidate involves *majority endorsement* if $k \geq (n-1)/2$, and it involves *super-majority endorsement* if $k \geq (n+1)/2$. Intuitively, (super-)majority endorsement means that the proposer and the players who are willing to endorse his proposal form a (super-)majority.

Proposition 5.5. *If a k-candidate with super-majority endorsement is an SSPE of the simplified ORBG, then it holds that $\delta(n+1) \leq 4$.*

The proof of Proposition 5.5 can be found in Appendix A. Since $n \geq 5$, one implication of this proposition is that a k-candidate with super-majority endorsement cannot be an SSPE if $\delta > 2/3$. Another implication is that, for any given $\delta > 0$, a k-candidate with super-majority endorsement cannot be an SSPE if the number of players satisfies $n > \frac{4}{\delta} - 1$, and thus if the number of players is not too small.

Proposition 5.6. *For any $\delta \in (0,1)$, there exists an odd integer n_δ sufficiently large, so that a k-candidate with majority endorsement cannot be an SSPE of the simplified ORBG if $n \geq n_\delta$.*

The proof of Proposition 5.6 can be found in Appendix A. Intuitively, the argument runs as follows: Consider an $((n-1)/2)$-candidate. Suppose that a proposer makes a unilateral deviation under which he offers one player only $\frac{\delta}{n} Y_k$ instead of δV_k. This player would no longer be willing to endorse the proposal. However, he would still be willing to vote in favor of the proposal once it was endorsed. In the formal proof of Proposition 5.6, we derive a parameter condition under which this deviation is profitable for the proposer, and we show that this condition boils down to an upper bound on n. In particular, the corollary below follows from the proof of Proposition 5.6.

Corollary 5.2. *A k-candidate with majority endorsement cannot be an SSPE of the simplified ORBG if δ is sufficiently close to one and if $n \geq 15$.*

The proof of Corollary 5.2 can be found in Appendix A.

Propositions 1 and 2 as well as Corollary 5.2 illustrate an important difference between our findings and those in Baron and Ferejohn (1989). They conclude that, with high n and high δ, equilibria involve majority endorsement by exactly $(n-1)/2$ players. This implies that, on the equilibrium path of play, the probability that the proposal on the floor is endorsed is equal to one half in each bargaining round. Since an endorsed proposal is always implemented in equilibrium, this also implies that the game ends in bargaining round t with probability $\frac{1}{2^{t+1}}$, which corresponds to an expected equilibrium delay of length one. Baron and Ferejohn's supposed equilibrium is based on strategies in which amendments are made in a more complicated way than with the "simple swaps" used here. We conclude that k-candidates with simple swaps that involve majority endorsement are not SSPE when n and δ are sufficiently high. In Section 8, we provide an example with $n = 51$ and δ close to one in which only 7 (rather than 25) of the 50 responding players endorse the proposal in equilibrium. In that example, the probability that any particular proposal is endorsed on the equilibrium path is only $7/50 = 0.14$ (instead of $1/2$). As a result, the expected length of equilibrium delay is more than six times as long as it would be with $k = 25$.[8]

So far, we have shown that for $n \geq 15$ and sufficiently large δ, a k-candidate with simple swaps can only be an SSPE if $k \leq (n-3)/2$. In that case, Eqns. (5.1)-(5.6) reduce to expressions that are continuous in δ. Hence, computing the limit behavior of the variables $V_k, W_k, X_k, Y_k, \lambda_k^+$, and λ_k^- when δ converges to one is equivalent to computing them while setting δ equal to one. Indeed, let us restate Eqns. (5.1)-(5.6) for $\delta = 1$ and $k \leq (n-3)/2$:

[8] On the path of play of a k-candidate with simple swaps, the probability that the proposal on the floor is endorsed (and then approved by majority voting) is $\frac{k}{n-1}$ in every round. Thus, the expected length of equilibrium delay can be written as $\frac{k}{n-1}\sum_{t=0}^{\infty}(1-\frac{k}{n-1})^t\, t = \frac{n-1-k}{k}$. For any n, if $k = \frac{n-1}{2}$, the expected length of delay is always one. In our example with $n = 51$ and $k = 7$, however, it is $\frac{51-1-7}{7} = \frac{43}{7} \approx 6.14$.

$$
\begin{aligned}
\overline{V}_k &= \left(\frac{k}{n-1}\right)\overline{X}_k + \left(\frac{n-1-k}{n-1}\right)\overline{W}_k, && (5.7)\\
\overline{W}_k &= (1-\overline{V}_k)/(n-1), && (5.8)\\
\overline{X}_k &= 1 - k\overline{V}_k - \left(\frac{n-1-2k}{2n}\right), && (5.9)\\
\overline{\lambda}_k^+ &= -\left(\frac{k+1}{n-1}\right)(\overline{V}_k - 1/n) + \left(\frac{1}{n-1}\right)(\overline{X}_k - \overline{W}_k), && (5.10)\\
\overline{\lambda}_n^- &= \left(\frac{k-1}{n-1}\right)(\overline{V}_k - 1/n) - \left(\frac{1}{n-1}\right)(\overline{X}_k - \overline{W}_k). && (5.11)
\end{aligned}
$$

Eqns. (5.7)-(5.9) are a system of three independent linear equations in three unknowns. We can solve this system for the variables $\overline{V}_k, \overline{W}_k$, and $\overline{X}_k$, and substitute the resulting expressions into Eqns. (5.10)-(5.11) to obtain:

$$
\begin{aligned}
\overline{\lambda}_k^+ &= (n-k-k^2)\left(\frac{n-1}{2n}\right)\left(\frac{1}{k^2(n-1)-k+n(n-1)}\right), && (5.12)\\
\overline{\lambda}_k^- &= (k^2-k-n)\left(\frac{n-1}{2n}\right)\left(\frac{1}{k^2(n-1)-k+n(n-1)}\right). && (5.13)
\end{aligned}
$$

Recalling that $n \geq 5$ and $1 \leq k \leq n-1$, it is easily verified that

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

Therefore, $\overline{\lambda}_k^+ > 0$ if and only if $n-k-k^2 > 0$, and $\overline{\lambda}_k^- > 0$ if and only if $k^2-k-n > 0$. Combined with Theorem 5.1, this implies Theorem 5.2.

Theorem 5.2. *Suppose that $n \geq 15$ and δ is sufficiently close to one. A k-candidate with simple swaps is an SSPE of the simplified ORBG if and only if the inequalities $k \leq (n-3)/2$ and $k^2 - k \leq n \leq k^2 + k$ are satisfied.*[9]

It follows that, for δ sufficiently close to one, the k-candidate equilibrium with simple swaps is an SSPE if $k \leq (n-3)/2$ and

$$k \in \left[-\frac{1}{2} + \sqrt{n+\frac{1}{4}},\ \ \frac{1}{2} + \sqrt{n+\frac{1}{4}}\right].$$

Corollary 5.3. *Suppose that $n \geq 15$ and δ is sufficiently close to one. There exists a unique $k = 1, \ldots, (n-3)/2$ such that the k-candidate with simple swaps is an*

[9] Recall that n is odd, and so the inequalities will always be strict.

k	$\overline{\lambda}_k^+$	$\overline{\lambda}_k^-$
1	6/115	0
2	-11/68	-3/85
3	-12/53	11/106
4	0	3/20

TABLE 5.1. THE CASE WITH $n = 5$ AND $\delta = 1$.

SSPE of the simplified ORBG. This k is the unique integer contained in the interval $\left[-\frac{1}{2} + \sqrt{n + \frac{1}{4}},\ \ \frac{1}{2} + \sqrt{n + \frac{1}{4}}\right]$.

The proof of Corollary 5.3 can be found in Appendix A.

Recall that we have assumed throughout the paper that n is odd and at least five. Theorem 5.2 and Corollary 5.3 do not say anything about the cases with $n \in \{5, 7, 9, 11, 13\}$ and δ sufficiently close to one. These cases can be dealt with by directly computing the relevant λ_k^+- and λ_k^--terms from Eqns. (5.1)-(5.6). For instance, Table 3 shows the results for $n = 5$. Indeed, we can see that $\overline{\lambda}_k^+$ and $\overline{\lambda}_k^-$ are both non-positive if and only if $k = 2$. When there are five players, then the 2-candidate with simple swaps is an SSPE.

5.8 Numerical illustration

5.8.1 Optimal choice of k

In this section, we illustrate our findings with some numerical examples. In Appendix B, we provide Mathematica code which can be used to replicate the numerical examples given in this paper (or to compute additional examples).

For $n = 51$ and various values of δ, Table 2 shows the unique value of k such that the k-candidate with simple swaps is an SSPE of the simplified ORBG . Recall that the payoffs induced by a k-candidate with simple swaps are given as the solutions to a system of equations which is continuous at $\delta = 1$. Therefore, we can find the limit values as δ converges to one by considering the relevant equations for $\delta = 1$.

In Proposition 5.4, we have shown that the $(n-1)$-candidate with simple swaps is an SSPE for sufficiently small δ. More precisely, in the proof of Proposition 5.4, we have derived the condition $\delta \leq \sqrt{\left(\frac{n-2}{2}\right)^2 + 1} - \left(\frac{n-2}{2}\right)$. For $n = 51$, we can compute

$$\sqrt{\left(\frac{n-2}{2}\right)^2 + 1} - \left(\frac{n-2}{2}\right) \approx 0.0204.$$

Indeed, the example illustrates the proposition: The 50-candidate with simple swaps is an SSPE for $\delta = 0.01$ and $\delta = 0.02$ but not for $\delta = 0.03$ and beyond. Moreover, we have shown in Proposition 5.5 that a k-candidate with simple swaps can only be an SSPE with super-majority endorsement if $\delta \leq 4/(n+1)$. If $n = 51$, then $4/(n+1) \approx 0.077$. Indeed, the example with $n = 51$ finds k-candidates with simple swaps to be SSPE for $\delta = 0.07$ and several lower values of δ. However, no such equilibria exist for $\delta = 0.08$ or above.

δ	Equilibrium value of k
0.01	50
0.02	50
0.03	41
0.04	35
0.05	32
0.06	29
0.07	27
0.08	25
0.09	24
0.1	23
$\vdots$	$\vdots$
0.2	16
$\vdots$	$\vdots$
0.5	10
$\vdots$	$\vdots$
1	7

TABLE 5.2. EQUILIBRIUM VALUES OF k FOR $n = 51$ AND VARIOUS VALUES OF δ.

We note that in all the numerical examples listed in the table, there is exactly one $k = 1, \ldots, n-1$, so that the k-candidate with simple swaps is an SSPE.

The example with $n = 51$ is also useful to illustrate the difference between our findings and those of Baron and Ferejohn (1989): They predict that the equilibrium number of endorsing players decreases when δ increases. However, they find that this number does not fall below $(n-1)/2 = 25$. Our Corollary 5.2 clearly contradicts this claim, and our calculations for the example with $n = 51$ illustrate that, indeed, the number of endorsing players continues to fall well below 25 as δ increases towards one. In particular, when δ is sufficiently close to one, the 7-candidate is an SSPE. In this case, only seven players are willing to endorse the equilibrium proposal.

As another example, let us consider the case where $n = 9$. Table 3 reports the equilibrium value of k for various specifications of δ.

If $n = 9$, then $\sqrt{\left(\frac{n-2}{2}\right)^2 + 1} - \left(\frac{n-2}{2}\right) \approx 0.14$, and, indeed, we can observe that the 8-candidate with simple swaps is an SSPE for $\delta = 0.1$, but not anymore for $\delta = 0.2$. We have shown that k-candidates with simple swaps and super-majority endorsement can only be SSPE if $\delta \leq 4/(n+1) = 0.4$. Indeed, for $\delta = 0.3$, we do find a k-candidate with simple swaps and super-majority endorsement that is an SSPE. For $\delta = 0.9$ and $\delta = 1$, the equilibrium value of k lies below $(n-1)/2$.

δ	Equilibrium value of k
0.1	8
0.2	6
0.3	5
0.4	4
0.5	4
0.6	4
0.7	4
0.8	4
0.9	3
1	3

Table 5.3. Equilibrium values for $n = 9$ and various values of δ.

5.8.2 Efficiency and equity for varying discount factors

One purpose of the original analysis by Baron and Ferejohn was to compare closed rule and open rule bargaining procedures with regard to the efficiency and equity of equilibrium outcomes. While open rules tend to lead to a more egalitarian distribution of the surplus, they suffer from inefficiencies. The reason is that the option of making amendments tends to lead to equilibrium delays, while closed rule bargaining models always predict immediate agreement. One important question is how one could weigh the efficiency loss against the equity gain.

We consider the example with $n = 51$ for different values of the discount factor. Examples 1 and 2 suggest that there is a large gain in fairness and no loss in efficiency when δ is either close to zero or close to one. However, Example 3 shows that for intermediate values of δ, the efficiency loss from open rules can be so large that even the responding players are ex ante better off than under a closed rule.

Example 1. Let us focus on the example where $n = 51$. First, we consider the case where δ is very small, say $\delta = 0.02$. If players were bargaining under a closed rule, agreement would be immediate and so the surplus divided would be of size one. The proposer would "buy" 25 players' votes by offering each of them the reservation payoff $\delta/n = 0.02/51 < 0.0004$. Hence, the proposer could secure a majority by offering less than $25 \times 0.0004 = 0.01$ to other players. Under closed rule, the proposer would be able to keep more than 99 percent of the surplus for himself.

Now we turn to the case of open rule bargaining. If $n = 51$ and $\delta = 0.02$, we have previously computed that $k = 50$. Since δ is so small, it is prohibitive for the proposer to risk bargaining delay. In equilibrium, agreement is reached immediately and the size of the surplus divided is one - just as it would be under closed rule bargaining. Substituting for n, δ, and k into Eqns. (5.1)-(5.4), we find that $V_{50} = X_{50} = 0.5$, while $W_{50} = 0.01$ and $Y_{50} = 1$. The equilibrium outcome under open rule can be described as follows: The proposer receives half of the surplus himself. He distributes the remaining half of the surplus equally to the other fifty players; hence, each of them receives $\delta V_{50} = 0.02 \times 0.5 = 0.01$, and is willing to endorse the proposal.

Clearly, the outcome under open rule is much more equitable than under closed rule, and equally efficient.

Example 2. Now we consider the example with $n = 51$ in the limit as $\delta \to 1$. First, suppose that bargaining takes place under the closed rule. In that case, the proposer needs to offer each of 25 players their reservation payoffs $\delta/n \to 1/51 \approx 0.0196$. The proposer can keep the remainder $1 - 25/51 \approx 0.5098$.

Now turn to the case of open rule bargaining. We have computed before that $n = 51$ and δ close to one give $k = 7$. By substitution into Eqns. (5.1)-(5.4), we find

$$\begin{aligned} V_7 &= \frac{7}{50} X_7 + \frac{43}{50} W_7, \\ W_7 &= (Y_7 - V_7)/50, \\ X_7 &= 1 - 7V_7 - \frac{18}{51}, \end{aligned}$$

where we observe that for δ close enough to one, the surplus converges to one. Solving this system we find

$$\begin{aligned} V_7 &\approx 0.054, \\ W_7 &\approx 0.0189, \\ X_7 &\approx 0.2693. \end{aligned}$$

Hence, under open rule, the proposer receives just over one fourth of the surplus (instead of more than a half under closed rule). Seven players receive 0.054 instead of just 0.0196. All other players obtain the same payoff under open rule bargaining as under closed rule bargaining. Indeed, the open rule bargaining leads to a more equitable, and equally efficient, outcome.

Example 3.

As an example, suppose that $n = 51$ and $\delta = 0.5$. With closed rule bargaining, 25 players would get the reservation payoff $\delta/n = 0.5/51 \approx 0.0098$. The proposer would keep the remaining $1 - 25 \times 0.0098 = 0.755$.

Under open rule bargaining, we find $k = 10$ and the relevant system of equations becomes

$$
\begin{aligned}
V_{10} &= 0.2X_{10} + 0.4W_{10}, && (5.14)\\
W_{10} &= (Y_{10} - V_{10})/50, && (5.15)\\
X_{10} &= 1 - 5V_{10} - \frac{15 \times 0.5}{51} Y_{10}, && (5.16)\\
Y_{10} &= 1/3. && (5.17)
\end{aligned}
$$

Solving this system yields $X_{10} \approx 0.4707$ and $\delta V_{10} \approx 0.048$. Hence, under open rule bargaining, one would expect the proposer to receive about 0.4707 and ten players to receive 0.048. Another 15 players would receive $0.5/51 \approx 0.0098$ and the remaining 25 players would receive nothing.

However, with open rule, the expected delay is 4 and the expected surplus is 1/3. So while the outcome under open rule is certainly more equitable than under closed rule, it is much less efficient.

Recall that V_{10} and W_{10} are the ex ante expected payoffs of the proposer and any player other than the proposer, respectively. From the above equations, we can compute $V_{10} \approx 0.096$ and $W_{10} = 0.0047$. In a closed rule bargaining game, the analogous ex ante payoffs would be 0.755 for the proposer (since agreement is immediate) and $0.5\frac{\delta}{n} = 0.0049$ for any other player. Note that ex ante, all players are better off with closed rule bargaining than with open rule bargaining for $\delta = 0.5$. The efficiency loss from delay is so great that even the gain in fairness cannot compensate the responders for it.

5.9 Return to the ORBG

In previous sections, we focused on the simplified ORBG. In the present section, we return to the original ORBG, as formally defined in Section 2. Recall the crucial difference between both games: In the original ORBG, whenever a player makes an amendment to a proposal on the floor, a vote determines whether or not the amendment replaces the proposal on the floor. We will show in this section that the main results and conclusions from our analysis of the simplified ORBG carry over to the ORBG itself. To this end, we first define a set of equilibrium candidates which are analogous to the k–candidates with simple swaps in the simplified ORBG. For this definition, we first recall the system of equations (5.1)-(5.4)

$$
\begin{aligned}
V_k &= \left(\frac{k}{n-1}\right) X_k + \left(\frac{n-1-k}{n-1}\right) \delta W_k, \\
W_k &= \frac{Y_k - V_k}{n-1}, \\
X_k &= 1 - k\delta V_k - \max\left\{0, \frac{n-1}{2} - k\right\} \left(\frac{\delta}{n}\right) Y_k, \\
Y_k &= \frac{k}{\delta k + (1-\delta)(n-1)}.
\end{aligned}
$$

Definition 5.2. *Consider the ORBG $G(\delta, n)$. For every $k = 1, \ldots, n-1$, a profile of stationary strategies is a* generalized k-candidate *if the following holds:*

1. *At every history $h \in H^{\emptyset}$, the proposer makes an anonymous proposal η which gives himself X_k, gives δV_k to k other players, gives $\frac{\delta}{n}((n-1)W_k + V_k)$ to $\max\{0, \frac{n-1}{2} - k\}$ more players, and zero to all remaining players.*
2. *At any history $h \in H_j^{i,\theta}$, player i expects the same payoff, say $p_{ij}(\theta, \sigma)$.*
3. *Let $p_i(\theta, \sigma) = \left(\frac{1}{n-1}\right) \sum_{j \in N \setminus \{i\}} p_{ij}(\theta, \sigma)$. For any $i_1, i_2 \in N$, it holds that $p_{i_2}(\pi_{i_1 \leftrightarrow i_2}(\theta), \sigma) = p_{i_1}(\theta, \sigma)$.*
4. *For every $\theta \in \Delta^n$, there is a set $T(\sigma, \theta) \subset \Delta^n$ such that the following holds: Whenever players vote between the proposal on the floor θ and some amendment θ', then the majority votes in favor of θ' if and only if $\theta' \in T(\sigma, \theta)$. Moreover, for every $\theta \in \Delta^n$, the set $T(\sigma, \theta)$ has the following properties: $\pi_{i \leftrightarrow j}(\theta) \in T(\sigma, \theta)$ and $\pi_{i \leftrightarrow j}(\widetilde{\theta}) \in T(\sigma, \pi_{i \leftrightarrow j}(\theta))$ if $\widetilde{\theta} \in T(\sigma, \theta)$ for any $i, j \in N$.*
5. *Whenever player i votes on an endorsed proposal θ, he votes in favor if and only if $\theta_i \geq \frac{\delta}{n}((n-1)W_k + V_k)$.*

Let us compare the generalized k–candidates to the k–candidates with simple swaps in Definition 5.1:

Points 1 and 5 in Definition 5.2 are familiar from Definition 5.1. In contrast to that earlier definition, however, we now also have to specify a *selection rule.* We impose on that selection rule a number of stationarity and anonymity requirements spelled out in Points 2, 3, and 4 of Definition 5.2. We verbally discuss these points in turn:

- Point 2 in Definition 5.2 imposes a *stationarity requirement* for histories at which an amendment can be made: Whenever player j can endorse or amend the same proposal θ made by the same player i, he acts in the same way. One implication is that player i has the same expected payoff, independently of the amendment, whenever his proposal θ is on the floor.

- Point 3 in Definition 5.2 adds an *anonymity requirement* to the previous point: The expected payoff of player i_1 when his proposal θ is on the floor is the same as the expected payoff of player i_2 when his proposal $\pi_{i_1 \leftrightarrow i_2}(\theta)$ is on the floor.

- Point 4 in Definition 5.2 puts *stationarity and anonymity restrictions* on the voting behavior when a proposal on the floor is pitted against an amendment: First, whenever the same proposal on the floor and the same amendment are pitted against each other, the winner is the same. Second, if the amendment is a simple swap of the proposal on the floor, then the amendment wins. Third, the majority's decision for an amendment or a proposal on the floor is unresponsive to a change in the players' "labels."

Consider a k–candidate with simple swaps which is an SSPE in the simplified ORBG. From that k–candidate with simple swaps, let us construct a generalized k–candidate by preserving the same *anonymous proposal,* the same *amendment rule,* and the same *voting rule* but adding the following *selection rule*: Whenever players choose between a proposal on the floor and an amendment, all players vote in favor of the amendment. It is trivially true that no unilateral deviation from this selection rule can improve a player's payoff. Hence, it is intuitive that the generalized k–candidate so constructed is an SSPE in the ORBG. This is formally stated in Proposition 5.7.

Proposition 5.7. *If the k-candidate with simple swaps is an SSPE of the simplified ORBG, then there is a generalized k-candidate with simple swaps that is an SSPE of the ORBG.*

The proof of Proposition 5.7 can be found in Appendix A.

Proposition 5.7 tells us that the SSPE found by studying the simplified ORBG corresponds to SSPE of the original ORBG. The remainder of this section asks whether the converse is also true: Can a generalized k–candidate be an SSPE of the ORBG without corresponding to a k–candidate with simple swaps that is an SSPE in the simplified ORBG? Loosely speaking, the issue in this section is what we have missed by restricting attention to the simplified ORBG. In order to address this question, we proceed in the following steps:

1. We consider the generalized $\widehat{k}$–candidate for some $\widehat{k}$. Suppose that it is optimal for the initial proposer to deviate unilaterally from the generalized $\widehat{k}$–candidate by making the proposal that he would make when playing according to the generalized k–candidate for some $k \neq \widehat{k}$. We show that, if this were indeed optimal, then any amendment to the proposal would be a simple swap of it. As a result, the initial proposer's unilateral one-shot deviation leads to a path of play that resembles the generalized k–candidate until a proposal is endorsed and voted upon. This is shown in Propositions 5.8 and 5.9 below.

2. We compute the payoff which the initial proposer could achieve by the aforementioned deviation. Due to the premise that this deviation is optimal for the initial proposer, it follows that the payoff we compute must be no less than the initial proposer's payoff from the generalized $\widehat{k}$–candidate. This gives us a necessary condition for the generalized $\widehat{k}$–candidate to be an SSPE. For the case with δ sufficiently close to one, we show that this necessary condition puts $\widehat{k}$ in a neighborhood around $\sqrt{n}$. This is Theorem 5.3 below.

3. Based on this insight, we argue that (except in a knife–edge case) there is only one $\widehat{k} \in \{1, \ldots, n-1\}$ so that the generalized $\widehat{k}$–candidate satisfies the necessary condition for an SSPE. (This is Corollary 5.4 below.)

4. Again for the case with δ close to one, we have already found a k such that the generalized k–candidate is an SSPE. We did so when we found an SSPE in the simplified ORBG and showed that it corresponded to one in the ORBG. As a result, the "uniqueness" that follows from Corollary 5.4 implies that our findings from the simplified ORBG carry over to the ORBG.

We introduce the following notation: Let σ be a generalized k-candidate. Take any sequence $(\theta^0, \theta^1, \ldots, \theta^T)$ such that with strictly positive probability, σ induces a path of play along which the proposal θ^0 is made at some history $h^0 \in H^\emptyset$ by player i^0. Then, the game reaches history h^1, where player i^1 makes the amendment θ^1, and it reaches history h^2, where player i^2 makes the amendment θ^2, and so on, until eventually player i^{T+1} endorses θ^T at history h^{T+1}.

Proposition 5.8. *Let $h \in H^\emptyset$ and $h' \in H^{i,\theta}$, for some $\theta \in \Delta^n$. Let player i be the proposer at h, and player j be the proposer at h'. Suppose that it is optimal for player i to propose θ at history h, provided that σ is played at all histories following h. Consider the choice of player j at history h'. Provided that σ is played at all histories following h', either it is optimal for player j to endorse the proposal θ at history h', or it is optimal to make the amendment $\pi_{i\leftrightarrow j}(\theta)$.*

Proof. Suppose that, at history h', it is strictly better for player j to make an amendment $\widetilde{\theta} \neq \pi_{i\leftrightarrow j}(\theta)$, instead of the amendment $\pi_{i\leftrightarrow j}(\theta)$. Thus $p_j(\widetilde{\theta}, \sigma) > p_j(\pi_{i\leftrightarrow j}(\theta), \sigma)$. By definition of a generalized k-candidate (see Definition 5.2), it follows that $p_i(\pi_{i\leftrightarrow j}(\widetilde{\theta}), \sigma) > p_i(\theta, \sigma)$. This implies that it is not optimal for player i to propose θ at history h, and the proof of the proposition is complete. $\square$

Repeating the same line of argument, we can also show the next proposition:

Proposition 5.9. *Let $h \in H^{(i,\theta)}$ and $h' \in H^{(j,\pi_{i\leftrightarrow j}(\theta))}$. Let player j be the proposer at h, and player k be the proposer at h'. Suppose that it is optimal for player j to make* *the amendment $\pi_{i\leftrightarrow j}(\theta)$. Consider the choice of player k at history h'. Provided that σ is played at all histories following h', either it is optimal for player k to endorse $\pi_{i\leftrightarrow j}(\theta)$, or it is optimal to make the amendment $\pi_{j\leftrightarrow k}(\pi_{i\leftrightarrow j}(\theta))$.*

The two propositions above lead to the following conclusions: Suppose that players $i^0, i^1, \ldots, i^T$ choose the initial proposal and the amendments optimally. Then, θ^t can be described as a simple swap of θ^{t-1} for any $t = 1, \ldots, T$. Moreover, if player i^0 at history h^0 has an expected payoff of V, then any player i^t who makes an amendment at history h^t with $t = 1, \ldots, T$, has an expected payoff of δV. Finally, player i^{T+1} endorses θ^T because that proposal gives him δV.

We compute the payoffs that would result if it were optimal for the initial proposer to deviate from the generalized $\widehat{k}$–candidate by making the anonymous proposal associated with the generalized k–candidate for some $k \neq \widehat{k}$. For this computation, we need the following system of equations:

$$
\begin{aligned}
V_k^{\widehat{k}} &= \left(\frac{k}{n-1}\right) X_k^{\widehat{k}} + \left(\frac{n-1-k}{n-1}\right) \delta W_k^{\widehat{k}}, && (5.18)\\
W_k^{\widehat{k}} &= \frac{Y_k - V_k^{\widehat{k}}}{n-1}, && (5.19)\\
X_k^{\widehat{k}} &= 1 - k\delta V_k^{\widehat{k}} - \max\left\{0, \frac{n-1}{2} - k\right\} \left(\frac{\delta}{n}\right) Y_{\widehat{k}}, && (5.20)\\
Y_{\widehat{k}} &= \frac{\widehat{k}}{\delta\widehat{k} + (1-\delta)(n-1)}, && (5.21)\\
Y_k &= \frac{k}{\delta k + (1-\delta)(n-1)}. && (5.22)
\end{aligned}
$$

Eqns. (5.18)–(5.20) are analogous to Eqns. (5.1)–(5.3). Given the premise that the initial proposer's deviation is optimal, we have shown in Propositions 5.8 and 5.9 that play proceeds according to the generalized k–candidate until a proposal is endorsed. Note that $\widehat{k}$ enters Eqns. (5.18)–(5.20) only through $\frac{\delta}{n} Y_{\widehat{k}}$, which is the continuation utility after some endorsed proposal has been voted on and rejected. In that case, play reverts back to the generalized $\widehat{k}$–candidate.[10]

From the above system of equations, we can compute the payoff $V_k^{\widehat{k}}$ for any pair $(k, \widehat{k})$.

[10] Note that the one-shot deviation principle is applied differently than in previous sections of this paper. In the present section, the choice between a proposal on the floor and an amendment is no longer rendered trivial: Therefore, it is no longer true that all histories at which a particular player is the proposer are "equivalent." It is true, however, that all histories in the set $H^{\emptyset}$ at which the same player proposes are followed by the same continuation game. Hence, we think of the open rule legislative bargaining game as a stochastic game. It moves to a new state whenever a history in $H^{\emptyset}$ is reached. If the proposer at that history is player i, then the game is in state i.

Suppose that, for some pair $(k, \widehat{k})$, we find $V_k^{\widehat{k}} \leq V_{\widehat{k}}^{\widehat{k}}$ for all $k \neq \widehat{k}$. Then the initial proposer's unilateral deviation which we have discussed above cannot be optimal. Clearly, this is a necessary condition for the generalized $\widehat{k}$–candidate to be an SSPE.

This is the gist of Theorem 5.3 below.

Theorem 5.3. *If the generalized $\widehat{k}$-candidate is an SSPE of the open rule legislative bargaining game $G(\delta, n)$, then $\widehat{k} \in \arg\max_{k \in \{1,\ldots,n-1\}} V_k^{\widehat{k}}$, where $V_k^{\widehat{k}}$ is the solution to Eqns. (5.18)-(5.22).*

Suppose that we wanted to test which of the generalized k-candidates satisfies the necessary condition established in Theorem 5.3. This would require the computation of the variables $(V_k^{\widehat{k}}, W_k^{\widehat{k}}, X_k^{\widehat{k}}, Y_{\widehat{k}}, Y_k)$ for $(n-1)^2$ possible pairs $(k, \widehat{k})$. Hence, testing for an SSPE is now considerably more complicated than it was in Section 6. Matters simplify a lot, however, when we focus on the case with patient players. Indeed, $Y_1, \ldots, Y_{n-1}$ all converge to one as $\delta \to 1$. As a result, $V^{\widehat{k}}, W^{\widehat{k}}$, and $X_k^{\widehat{k}}$ converge to limits that are independent of $\widehat{k}$. This allows us to simplify the Eqns. (5.18)-(5.20). For each $k = 1, \ldots, n-1$, let $\overline{V}_k, \overline{W}_k$, and $\overline{X}_k$ be the solution to the following system of equations:

$$\overline{V}_k = \left(\frac{k}{n-1}\right)\overline{X}_k + \left(\frac{n-1-k}{n-1}\right)\overline{W}_k, \tag{5.23}$$

$$\overline{W}_k = (1-\overline{V}_k)/(n-1), \tag{5.24}$$

$$\overline{X}_k = 1 - k\overline{V}_k - \max\left\{0, \frac{n-1}{2} - k\right\}\left(\frac{1}{n}\right). \tag{5.25}$$

Note that Eqns. (5.23)-(5.25) are the same system of linear equations as in Section 5, now applied at $\delta = 1$.

Now we are ready to state the main result of this section.

Theorem 5.4. *Suppose that $n \geq 15$ and δ is sufficiently close to one. If a generalized k-candidate is an SSPE of the ORBG, then it holds that*

$$k \in \{1, \ldots, n-1\} \cap (\sqrt{n} - 1, \sqrt{n} + 1).$$

The proof of Theorem 5.4 is relegated to Appendix A.

In order to assess the implications of Theorem 5.4, let us first consider the case where n is such that $\sqrt{n}$ is an integer. In that case, $\sqrt{n}$ is the only integer contained in the interval $(\sqrt{n} - 1, \sqrt{n} + 1)$. Hence, the generalized $\sqrt{n}$-candidate is the only generalized k-candidate that can be an SSPE.

Now consider the case where n is such that $\sqrt{n}$ is not an integer. In that case, the interval $(\sqrt{n} - 1, \sqrt{n} + 1)$ contains two integers. Let us denote them by k^* and $k^* + 1$.

Now we use Eqns. (5.23)-(5.25) to compute $\overline{V}_{k^*}$ and $\overline{V}_{k^*+1}$. There is no reason to expect that these two amounts are generally equal. If $\overline{V}_{k^*} > \overline{V}_{k^*+1}$, then Theorem 5.3 implies that the generalized (k^*+1)-candidate cannot be an SSPE. Similarly, if $\overline{V}_{k^*} < \overline{V}_{k^*+1}$, then Theorem 5.3 implies that the generalized k^*-candidate cannot be an SSPE.

From these observations, we obtain the following corollary:

Corollary 5.4. *Suppose that $n \geq 15$ and δ is sufficiently close to one. Suppose that the k^*-candidate with simple swaps is an SSPE of the ORBG. Moreover, suppose that there is some $k^{**} \neq k^*$ such that the generalized k^{**}-candidate is an SSPE. Then, k^* and k^{**} are successive integers and it holds that $\overline{V}_{k^*} = \overline{V}_{k^{**}}$.*

The interpretation is as follows: Consider the case where $n \geq 15$ and δ is close enough to one. In our analysis of the simplified ORBG, we have found one k such that the k–candidate with simple swaps is an SSPE. In Proposition 7, we have shown that this SSPE of the simplified ORBG easily extends to an SSPE of the ORBG. Hence, we already have found one particular k such that the generalized k–candidate is an SSPE of the ORBG. Theorem 5.4 and Corollary 5.4 tell us that the SSPE which we have already found is (except in a knife–edge case) the *only* generalized k–candidate that is an SSPE. In addition, even in that knife–edge case, there can at most be two values of k such that the generalized k–candidate that is an SSPE, these two values of k must be successive integers, and the proposer's payoffs in both potential equilibria are equal. The conclusion is that our main results from the simplified ORBG carry over to the ORBG.

5.10 Conclusion

While open rule legislative bargaining is common in reality, the bargaining literature has focused on closed rule procedures to a large extent. In contrast to this literature, we reconsidered the leading model of open rule legislative bargaining. While it is true that the analysis of open rule procedures involves many more complications than a closed rule procedure, it is possible to derive some conclusions that allow for a comparison between the two bargaining procedures. In particular, we have devised a method to construct and test equilibrium candidates for any values of underlying model parameters, in particular the size of the legislature and the discount factor. One important insight is that when the legislature is large, or its members are patient, the probability that bargaining leads to an agreement without delay falls below one half. Hence, substantial delays are to be expected in equilibrium when bargaining takes place under an open rule. The advantage of an open rule is, however, that it tends to lead to more egalitarian surplus divisions than a closed rule.

One contribution of our paper is to study a class of stationary equilibria that are more tractable than those suggested by Baron and Ferejohn. Our analysis also qualifies some of the conclusions reached in existing work on the open rule legislative bargaining model. While we confirm Baron and Ferejohn's insight that open rules tend to lead to less efficient and more egalitarian outcomes, we find that they may be even less efficient and also less egalitarian than suggested by them.

Moreover, our analysis may have important implications for the design of legislatures and their committees. For instance, the tendency of open rules to produce egalitarian outcomes, even after the proposer has been selected at the cost of significant delays, opens up a more detailed comparison of the egalitarian efficiency trade-offs between closed and open rules. Smaller legislatures yield less delay and a more egalitarian allocation than larger legislatures under open rules.

When the size of the legislature may need to be quite large for other reasons than examined in this paper, e.g. to be sufficiently representative of the underlying electorate, surplus division could be first delegated to a smaller committee that itself is representative of the legislature. If the committee uses the open rule and the committee decision is put to a final vote in the legislature under a closed rule, the efficiency and equality advantages of open rules could be preserved.

Another finding is that patient players induce more delays, even to the extreme that the expected delay becomes arbitrarily large. Hence, from an efficiency perspective, it would be useful if players are more impatient when deciding about surplus divisions than they actually are. This could be achieved by limiting the time members of a legislature can spend on committees to take or at least to prepare decisions on surplus divisions. These and other possible applications and extensions should be pursued in future work.

Appendix A

Proof of Proposition 5.1. The proof consists of three steps. First, we show part (i) and part (ii) for the case $k \geq (n-1)/2$. Second, we show part (ii) for the complementary case $k \leq (n-3)/2$. Third, we show part (iii).

Step 1. Suppose first that $k \geq (n-1)/2$. From a direct computation of the solution of Eqns. (5.1)-(5.4), we find

$$\begin{aligned} \beta_k V_k/k &= (n-1)^2 - \delta(n-2)(n-1-k), \\ \beta_k W_k/k &= \delta(n-1-k+k^2), \end{aligned}$$

where β_k is given by

$$\beta_k = [(1-\delta)(n-1)+\delta k][(n-1)^2 + \delta(n-1)(1+k^2) - \delta k].$$

We observe that β_k is strictly positive since $(n-1)^2 > \delta k$. Hence, the difference $V_k - W_k$ has the same sign as $(\beta_k/k)(V_k - W_k)$. The latter expression can be written as

$$\begin{aligned} (V_k - W_k)(\beta_k/k) &= (n-1)^2 - \delta(n-2)(n-1-k) - \delta(n-1-k+k^2) \\ &= (1-\delta)(n-1)^2 - \delta k^2 + \delta k(n-1) \\ &= (1-\delta)(n-1)^2 + \delta k(n-1-k). \end{aligned}$$

We see that $V_k = W_k > 0$ if $\delta = 1$ and $k = n-1$. If at least one of the inequalities $\delta \leq 1$ and $k \leq n-1$ holds strictly, we see that $V_k > W_k > 0$.

Step 2. Suppose now that $k \leq (n-3)/2$. In that case, from Eqns. (5.1)-(5.4), we can directly compute

$$\begin{aligned} \gamma_k V_k/k &= 2n(n-1)^2 - \delta[2n(n^2-3n+2) - k(n^2-2n-1) - 2k^2(n-1)], \\ \gamma_k W_k/k &= \delta[2n(n-1) - (n+1)k + 2k^2(n-1)], \end{aligned}$$

where γ_k is given by

$$\gamma_k = 2n[(1-\delta)(n-1)+\delta k][(n-1)^2 + \delta(n-1)(1+k^2) - \delta k].$$

Since $\gamma_k > 0$, we can conclude that $V_k - W_k$ has the same sign as the expression

$$\begin{aligned} (\gamma_k/k)(V_k - W_k) &= 2n(n-1)^2 - \delta[2n(n-1)^2 - k(n^2-n)] \\ &= 2n(1-\delta)(n-1)^2 + \delta k(n^2-n) > 0. \end{aligned}$$

Now it remains to show that $W_k > 0$. To this end, Note that

$$\begin{aligned}\gamma_k W_k/(k\delta) &= 2n(n-1) - (n+1)k + 2(n-1)k^2 \\ &= k^2(2n-2) - k(n+1) + (2n-2)n.\end{aligned}$$

Since $n+1 < 2n-2$ and $k < n$, this expression is strictly positive, and hence $W_k > 0$, as desired.

Step 3.

We have already shown that, for any $k = 1, \ldots, n-1$, it holds that $W_k > 0$ and $V_k - W_k \geq 0$, implying in particular that $V_k > 0$ and $V_k - \delta W_k \geq 0$. From the explicit Eqn. (5.4), it follows directly that $Y_k > 0$. Finally, Eqn. (5.1) can be rewritten as

$$kX_k = (n-1)(V_k - \delta W_k) + k\delta W_k,$$

implying that $X_k > 0$ as well.

$\square$

Proof of Corollary 5.1.

Part (i) of the corollary follows directly from substituting $k = n-1$ and $\delta = 1$ into Eqn. (5.1). In order to see why part (ii) of the corollary is true, recall from Proposition 5.1 that for any choices of k and δ other than $k = n-1$ and $\delta = 1$, we have $V_k > W_k$. Thus, Eqn. (5.1) implies the inequality $(n-1)V_k < kX_k + (n-1-k)\delta V_k$. Rearranging yields

$$X_k\left(\frac{k}{(1-\delta)(n-1)+\delta k}\right) > V_k.$$

It only remains to note that $\frac{k}{(1-\delta)(n-1)+\delta k} \leq 1$, since $k \leq n-1$ and the denominator is larger than k. Hence, we can conclude that $X_k > V_k \geq \delta V_k$, as desired.

$\square$

Proof of Proposition 5.3.

Suppose that there is a vector $\widehat{\theta} \in \Delta^n$ such that proposing $\widehat{\theta}$ instead of the proposal prescribed by the k-candidate with simple swaps is a profitable deviation for the proposer, say player i. Suppose that there is a player $j \in N\backslash\{i\}$ such that $0 < \widehat{\theta}_j < \frac{\delta}{n}((n-1)W_k + V_k)$. Player j neither endorses the proposal $\widehat{\theta}$, nor does he vote in its favor. Consequently, it would also be a profitable deviation for player i to offer zero to player j, and offer $\widehat{\theta}_l$ to all players $l \in N \setminus \{j\}$. By the same token, suppose that there is a player $j \in N \setminus \{i\}$ such that $\frac{\delta}{n}((n-1)W_k + V_k) < \widehat{\theta}_j < \delta V_k$. In that case, player j is willing to vote in favor of $\widehat{\theta}$, but not willing to endorse it. This would not change if he was offered $\frac{\delta}{n}((n-1)W_k + V_k)$ instead of $\widehat{\theta}_j$. Thus, it would also be a profitable deviation for player i to offer player j only $\frac{\delta}{n}((n-1)W_k + V_k)$, and offer each player $l \in N \setminus \{j\}$ the amount $\widehat{\theta}_l$. Repeating the same argument, we see that if the proposer has any profitable deviation $\widehat{\theta}$, then he has a profitable deviation to a proposal $\widetilde{\theta}$ which gives each player other than the proposer either zero, or $\frac{\delta}{n}((n-1)W_k + V_k)$, or δV_k.

Let $\mathcal{P}_k$ be the set of vectors $\theta \in \Delta^n$ that, for some $k' \in \{1, \ldots, n-1\}$, contain k' components equal to δV_k and $\max\{0, \frac{n-1}{2} - k'\}$ components equal to $\frac{\delta}{n} Y_k$, where V_k and Y_k are as defined in Eqns. (5.1)-(5.4). Moreover, a vector $\theta \in \mathcal{P}_k$ has one component equal to $1 - k'\delta V_k - \max\{0, \frac{n-1}{2} - k'\}\frac{\delta}{n} Y_k$. Any remaining components are equal to zero.[11]

Let λ_k^{+m} be the gain which the proposer can make by offering δV_k to $k+m$ players instead of to k players, starting from a proposal $\theta \in \mathcal{P}_k$, where $m \geq 2$. We want to show that $\lambda_k^{+m} > 0$ implies $\lambda_k^{+} > 0$. Suppose that $k \geq (n-1)/2$. Then we have

$$\lambda_k^{+m} = -\left(\frac{k}{n-1}\right) m\delta V_k + \left(\frac{m}{n-1}\right)(X_k - \delta W_k - m\delta V_k).$$

If this is strictly positive, then dividing by m yields

$$-\left(\frac{k}{n-1}\right)\delta V_k + \left(\frac{1}{n-1}\right)(X_k - \delta W_k) - \left(\frac{m}{n-1}\right)\delta V_k > 0,$$

which can be rewritten equivalently as

$$-\left(\frac{k}{n-1}\right)\delta V_k + \left(\frac{1}{n-1}\right)(X_k - \delta W_k - \delta V_k) > \left(\frac{m-1}{n-1}\right)\delta V_k,$$

hence

$$\lambda_k^{+} > \left(\frac{m-1}{n-1}\right)\delta V_k.$$

We have shown earlier that $V_k > 0$. Thus, it follows that $\lambda_k^{+} > 0$, as desired.

Now consider the case where $k \leq (n-3)/2$. Then, we have

$$\lambda_k^{+m} \leq -\left(\frac{k}{n-1}\right) m\delta V_k + \left(\frac{m}{n-1}\right)(X_k - \delta W_k - m\delta V_k).$$

Suppose that $\lambda_k^{+m} > 0$, then dividing by $m/(n-1)$ and rearranging terms, it follows that

$$-(k+m)\delta V_k + X_k - \delta W_k > 0.$$

Since $m \geq 2$, we can conclude that also

$$-(k+1)\delta V_k + X_k - \delta W_k > 0.$$

[11] Recall that the set Δ^n consists of vectors that are non-negative in all components and sum up to (at most) one. It follows that the number k' satisfies the inequality $1 - k'\delta V_k - \max\{0, \frac{n-1}{2} - k'\}\frac{\delta}{n} Y_k \geq 0$. In other words, it is ensured that for any proposal in $\mathcal{P}_k$, the share of surplus for the proposer remains non-negative. We also show in Corollary 5.1 that the proposer's share is larger than that of any other player.

Again, a fortiori, we have that

$$-(k+1)\left(\delta V_k - \frac{\delta}{n} Y_k\right) + X_k - \delta W_k > 0,$$

and hence $\lambda_k^+ > 0$, as desired.

An analogous argument can be used to show that $\lambda_k^{-m} > 0$ implies $\lambda_k^- > 0$. $\square$

Proof of Proposition 5.4. For $k = n-1$, the system of Eqns. (1)-(4) yields the following solutions:

$$\begin{aligned} V_{n-1} &= \frac{1}{1+\delta(n-1)}, \\ W_{n-1} &= \frac{\delta}{1+\delta(n-1)}, \\ X_{n-1} &= \frac{1}{1+\delta(n-1)}, \\ Y_{n-1} &= 1. \end{aligned}$$

By definition, $\lambda_{n-1}^+ = 0$, so it remains to show that for δ sufficiently small, $\lambda_{n-1}^- \leq 0$. Substituting the above equations, we obtain

$$\lambda_{n-1}^- = \frac{\delta(n-2) - 1 + \delta^2}{(n-1)(1+\delta(n-1))}.$$

Indeed, it follows that $\lambda_{n-1}^- \leq 0$ if and only if $\delta^2 + \delta(n-2) - 1 \leq 0$. This inequality is satisfied when

$$\delta \leq \sqrt{\left(\frac{n-2}{2}\right)^2 + 1} - \left(\frac{n-2}{2}\right).$$

$\square$

Proof of Proposition 5.5. Suppose that there is a $k \geq (n+1)/2$ such that some k-candidate is an SSPE. Now consider a deviation by the initial proposer from the supposed SSPE. Under the deviation, the proposer offers δV_k to $k-1$ instead of to k players, and offers zero to the remaining $n-k$ players, where V_k is as before. This deviation gives the proposer an expected payoff of

$$\left(\frac{k-1}{n-1}\right)(1 - k\delta V_k + \delta V_k) + \left(\frac{n-k}{n-1}\right)\delta W_k,$$

while the proposer's expected payoff when playing according to the supposed SSPE is

$$V_k = \left(\frac{k}{n-1}\right)(1 - k\delta V_k) + \left(\frac{n-1-k}{n-1}\right)\delta W_k.$$

Clearly, a necessary condition for the k-candidate with simple swaps to be an SSPE is the inequality

$$\left(\frac{k-1}{n-1}\right)\delta V_k - \left(\frac{1}{n-1}\right)(1-k\delta V_k) + \delta\left(\frac{1}{n-1}\right)W_k \leq 0.$$

Since $\delta\left(\frac{1}{n-1}\right)W_k \geq 0$, it is necessary that

$$\left(\frac{k-1}{n-1}\right)\delta V_k - \left(\frac{1}{n-1}\right)(1-k\delta V_k) \leq 0. \tag{5.26}$$

The inequality can be rearranged to

$$\delta V_k \leq \frac{1}{2k-1}.$$

It follows that

$$1 - k\delta V_k \geq \frac{k-1}{2k-1}.$$

Recall that V_k is the expected payoff of the proposer induced by the supposed SSPE. Under that strategy profile, the proposer offers himself $1 - k\delta V_k$, and the proposal is endorsed (and then implemented) with probability $k/(n-1)$. Thus, we obtain

$$V_k \geq \left(\frac{k}{n-1}\right)(1-k\delta V_k) \geq \left(\frac{k}{n-1}\right)\left(\frac{k-1}{2k-1}\right).$$

Combining the above inequalities, we obtain

$$\frac{1/\delta}{2k-1} \geq V_k \geq \left(\frac{k}{n-1}\right)\left(\frac{k-1}{2k-1}\right).$$

This leads to the condition

$$\delta \leq \frac{n-1}{k^2-k} \leq \frac{4(n-1)}{(n-1)(n+1)} = \frac{4}{n+1}.$$

The last inequality follows from the premise that $k \geq (n+1)/2$. Canceling $(n-1)$, we obtain $\delta(n+1) \leq 4$, as desired. □

Proof of Proposition 5.6. In view of Proposition 5.5, we only have to consider the case where $k = (n-1)/2$. Indeed, fix some value of $\delta \in (0,1)$ and a number n of players, and suppose that the $\frac{n-1}{2}$-candidate with simple swaps is an SSPE. From these premises, we are going to derive an implicit upper bound on n. On the path of play of the supposed SSPE, every proposal is endorsed with probability $1/2$, and is accepted with certainty once it is endorsed. Thus, we find the following expression for the expected size of the total surplus divided:

$$V_k + (n-1)W_k = \left(\frac{1}{2}\right)\sum_{t=0}^{\infty}\left(\frac{\delta}{2}\right)^t = 1/(2-\delta).$$

Due to sincere voting (Point 3 in Definition 1), a player accepts an endorsed proposal if and only if it gives him at least $\left(\frac{\delta}{n}\right)\left(\frac{1}{2-\delta}\right)$. Consider a deviation from the supposed SSPE by the current proposer. This deviation consists of changing the offer to one player from δV_k to $\left(\frac{\delta}{n}\right)\left(\frac{1}{2-\delta}\right)$. Consequently, that player will no longer endorse the proposal but will still vote for it once it has been endorsed by some other player.

This deviation gives the proposer an expected payoff of

$$\left(\frac{n-3}{2(n-1)}\right)\left(X_k + \delta V_k - \frac{\delta}{n}\frac{1}{2-\delta}\right) + \left(\frac{1}{2} + \frac{1}{n-1}\right)\delta W_k.$$

Recall that playing according to the supposed SSPE gives a proposer an expected payoff of

$$V_k = \frac{1}{2}X_k + \frac{1}{2}\delta W_k.$$

Consequently, the proposer's gain from the deviation can be written as

$$\left(\frac{n-3}{2(n-1)}\right)\left(\delta V_k - \left(\frac{\delta}{n}\right)\left(\frac{1}{2-\delta}\right)\right) + \left(\frac{1}{n-1}\right)\delta W_k - \left(\frac{1}{n-1}\right)X_k.$$

Taking into account that $X_k = 1 - \frac{n-1}{2}\delta V_k$, we find the following necessary condition for the $\frac{n-1}{2}$-candidate with simple swaps to be an SSPE:

$$\left(\frac{n-3}{2(n-1)}\right)\left(\delta V_k - \left(\frac{\delta}{n}\right)\left(\frac{1}{2-\delta}\right)\right) - \left(\frac{1}{n-1}\right)\left(1 - \left(\frac{n-1}{2}\right)\delta V_k\right) + \left(\frac{1}{n-1}\right)\delta W_k \leq 0.$$

In view of the fact that $W_k \geq 0$, this implies the necessary condition

$$\left(\frac{n-3}{2(n-1)}\right)\left(\delta V_k - \frac{\delta}{n}\ \frac{1}{2-\delta}\right) - \left(\frac{1}{n-1}\right)\left(1 - \left(\frac{n-1}{2}\right)\delta V_k\right) \leq 0.$$

This yields

$$\delta V_k \leq \frac{2n(2-\delta) + \delta(n-3)}{2n(2-\delta)(n-2)}.$$

In the supposed SSPE, any proposal is endorsed with probability 1/2. If a player's proposal is endorsed, then it is also accepted, and so the expected payoff of a proposer can be bounded as follows:

$$V_k \geq \left(\frac{1}{2}\right)\left(1 - \left(\frac{n-1}{2}\right)\delta V_k\right).$$

Using the bound previously derived for δV_k, we can write

$$V_k \geq \frac{1}{2} - \left(\frac{n-1}{4}\right)\left(\frac{2n(2-\delta)+\delta(n-3)}{2n(2-\delta)(n-2)}\right).$$

Now we have bounded V_k both from above and below, as follows:

$$\left(\frac{1}{\delta}\right)\left(\frac{2n(2-\delta)+\delta(n-3)}{2n(2-\delta)(n-2)}\right) \geq V_k \geq \frac{1}{2} - \left(\frac{n-1}{4}\right)\left(\frac{2n(2-\delta)+\delta(n-3)}{2n(2-\delta)(n-2)}\right).$$

This readily implies the inequality

$$\left(\frac{2n(2-\delta)+\delta(n-3)}{2n(2-\delta)(n-2)}\right)\left(\frac{1}{\delta}+\frac{n-1}{4}\right) \geq 1/2.$$

Arranging this inequality in the quadratic form yields

$$\delta^2\left(3n^2-10n+3\right) - \delta\left(4n^2-8n+12\right) + 16n \geq 0. \tag{5.27}$$

Recall that we have assumed throughout the paper that $n \geq 5$, which implies that $n^2 - 2n > 0$ and $3n^2 - 10n + 3 > 0$. Therefore, it follows that

$$\delta^2 - \delta\left(\frac{4n^2-8n+12}{3n^2-10n+3}\right) + \left(\frac{16n}{3n^2-10n+3}\right) \geq 0. \tag{5.28}$$

It is easily verified that for any $n \geq 13$,[12] we have

$$\left(\frac{4n^2-8n+12}{6n^2-20n+6}\right)^2 - \left(\frac{16n}{3n^2-10n+3}\right) > 0, \tag{5.29}$$

and consequently, the quadratic Inequality (5.27) has two distinct real roots

$$\overline{\delta}(n) = \frac{4n^2-8n+12}{6n^2-20n+6} + \sqrt{\left(\frac{4n^2-8n+12}{6n^2-20n+6}\right)^2 - \left(\frac{16n}{3n^2-10n+3}\right)}, \tag{5.30}$$

$$\underline{\delta}(n) = \frac{4n^2-8n+12}{6n^2-20n+6} - \sqrt{\left(\frac{4n^2-8n+12}{6n^2-20n+6}\right)^2 - \left(\frac{16n}{3n^2-10n+3}\right)}. \tag{5.31}$$

Existence of the supposed SSPE requires that either $0 < \delta \leq \underline{\delta}(n)$ or $\overline{\delta}(n) \leq \delta \leq 1$. However, in the limit, as $n \to \infty$, we find that $\overline{\delta}(n)$ converges to $4/3 > 1$, while $\underline{\delta}(n)$ converges to zero. This implies that Inequality (5.27) cannot be satisfied for n sufficiently large. In turn, this implies that a k-candidate with $k \geq (n-1)/2$ ("majority endorsement") is not an SSPE if n is sufficiently large. $\square$

[12]In fact, if we consider the left-hand side of Ineq. (5.29) as a continuous function of n, then we find that it is equal to zero for $n \approx 3.52$ and $n \approx 11.99$, negative for values of n between these two roots, and positive otherwise. Recall that in our model, we assume that n is an odd integer and that $n \geq 5$.

Proof of Corollary 5.2.

We have shown in Proposition 5.5 that a k–candidate with super–majority endorsement cannot be an SSPE if δ is sufficiently close to one. Hence, the have to show the claim of the corollary only in the special case with $k = (n-1)/2$. In that case, we have shown in the proof of Proposition 5.6, that Ineq. (28) is a necessary condition for the $(n-1)/2$–candidate with simple swaps to be an SSPE. Observe that this inequality changes continuously with δ, and that it reduces to $n^2 - 14n + 9 \leq 0$ as δ converges to one. It is easily verified that this inequality is violated for any $n > 7 + 2\sqrt{10} \approx 13.32$. Since n is assumed to be odd, the corollary follows.

□

Proof of Corollary 5.3.

We need to show that the interval $\left[-\frac{1}{2} + \sqrt{n + \frac{1}{4}},\ \ \frac{1}{2} + \sqrt{n + \frac{1}{4}}\right]$ contains exactly one integer. Since this interval is closed and of unit length, it can at most contain two integers. It is sufficient to show that it does not contain two integers. Indeed, suppose towards a contradiction that the endpoints of the interval are integers. In particular, let $z = -\frac{1}{2} + \sqrt{n + \frac{1}{4}}$ be an integer. Solving for n yields $n + \frac{1}{4} = (z + \frac{1}{2})^2$. This can be rewritten as $n + \frac{1}{4} = z^2 + z + \frac{1}{4}$ or as $n = z(z+1)$. Either z is even, or $z+1$ is even. Therefore, the product $z(z+1)$ is even, and so n is even. This contradicts our assumption that n is odd.

□

Proof of Proposition 5.7

The proof of this proposition consists of two steps:

We construct a generalized k-candidate in the following way: Suppose that the stationary strategy profile σ^* is a k-candidate with simple swaps and that it is an SSPE of the simplified ORBG. The profile σ^* consists of anonymous proposals η^{i*} for every $i \in N$, amendment rules ψ^{i*} for every $i \in N$, and acceptance rules A^{i*} for every $i \in N$. Let σ^{**} be a stationary strategy profile for the ORBG, which consists of these same anonymous proposals η^{i*} for every $i \in N$, amendment rules ψ^{i*} for every $i \in N$, and acceptance rules A^{i*} for every $i \in N$, and, in addition, features the following selection rule χ^{**}, used by every player $i \in N$: *"Whenever a vote takes place between an amendment and a proposal on the floor, every player votes in favor of the amendment."* It is trivially true that a unilateral deviation by any player from the selection rule χ^{**} does not change the outcome of a vote, and therefore cannot be profitable for the deviating player. If there were profitable unilateral deviations at histories other than the ones governed by the selection rule, then the k-candidate with simple swaps would not be an SSPE of the simplified ORBG. Hence, we have shown that the stationary strategy profile σ^{**} is an SSPE of the ORBG.

Finally, we observe that the stationary strategy profile σ^{**} is a generalized k-candidate: The anonymous proposal and the acceptance rule under σ^{**} satisfy Points 1 and 5, respectively, of Definition 5.2. Note that they correspond exactly to Points 1 and 3 of Definition 5.1. Under σ^{**}, every player always votes in favor of any amendment against the proposal on the floor, which trivially satisfies Point 4 of Definition 5.2. The amendment rule under σ^{**} corresponds to that in a k-candidate with simple swaps as defined in Point 2 of Definition 5.1. This amendment rule, combined with the trivial selection rule, ensures that Points 2, 3, and 4 of Definition 5.2 are satisfied.

$\square$

Proof of Theorem 5.4.

We solve Eqns. (5.23)-(5.25) for $\overline{V}_k$:

$$\overline{V}_k = \begin{cases} \frac{2k^2(n-1)+2(n-1)n+k(n^2-2n-1)}{2n(k^2(n-1)+n(n-1)-k)}, & \text{if } k \le (n-3)/2, \\ \frac{n+k(n-2)-1}{k^2(n-1)+n(n-1)-k}, & \text{if } k \ge (n-1)/2. \end{cases} \tag{5.32}$$

We have already argued that the generalized $\widehat{k}$-candidate can only be an SSPE if $\overline{V}_{\widehat{k}} \ge \overline{V}_k$ for every $k = 1, \dots, n-1$.

As an auxiliary, it is useful to define the continuous function

$$\nu(\kappa) := \frac{2\kappa^2(n-1) + 2(n-1)n + \kappa(n^2 - 2n - 1)}{2n(\kappa^2(n-1) + n(n-1) - \kappa)}$$

for the real-valued variable κ. We observe that $\nu(\kappa)$ exists for all $\kappa \in [1, (n-3)/2]$. Its derivative can be written as

$$\nu'(\kappa) = -\frac{(n-1)^3(\kappa^2 - n)}{2n(\kappa^2(n-1) - \kappa + n(n-1))},$$

and we can easily verify that

$$\begin{aligned} \nu'(\kappa) &> 0 \text{ if } \kappa \in [1, \sqrt{n}), \\ \nu'(\kappa) &= 0 \text{ if } \kappa = \sqrt{n}, \\ \nu'(\kappa) &< 0 \text{ if } \kappa \in (\sqrt{n}, (n-3)/2]. \end{aligned}$$

We see that $\nu(\kappa)$ attains its unique maximum at the point $\kappa = \sqrt{n}$. So far, we have shown that the generalized $\widehat{k}$-candidate can only be an SSPE if either $\sqrt{n} - 1 < \widehat{k} < \sqrt{n} + 1$ or $\widehat{k} \ge \frac{n-1}{2}$. In order to complete the proof of the theorem, we want to exclude the latter possibility.

To this end, we first note that $\overline{V}_{\widehat{k}} \le \nu(\widehat{k})$ for any $\widehat{k} \ge \frac{n-1}{2}$. In order to see this, consider the following sequence of implications:

$$
\begin{aligned}
\widehat{k} &\geq \frac{n-1}{2}, \\
0 &\leq 2\widehat{k} - (n-1), \\
0 &\leq \widehat{k}^2 \left(\frac{n-1}{n}\right) - \left(\frac{\widehat{k}}{2}\right)\left(\frac{(n-1)^2}{n}\right), \\
0 &\leq \widehat{k}^2 \left(\frac{n-1}{n}\right) + \widehat{k}\left(\frac{n^2-2n-1}{2n}\right) - \widehat{k}(n-2), \\
n + \widehat{k}(n-2) - 1 &\leq \widehat{k}^2 \left(\frac{n-1}{n}\right) + (n-1) + \frac{\widehat{k}}{2n}\left(n^2 - 2n - 1\right), \\
\overline{V}_{\widehat{k}} &\leq \nu(\widehat{k}).
\end{aligned}
$$

Now suppose by way of contradiction that for some $\widehat{k} \geq \frac{n-1}{2}$, the generalized $\widehat{k}$-candidate is an SSPE. Then, for any $k \in \{1, \ldots n-1\} \setminus \{\widehat{k}\}$ it holds that

$$\overline{V}_k \leq \overline{V}_{\widehat{k}} \leq \nu(\widehat{k}).$$

Due to the premise that $n \geq 15$, we have that $\sqrt{n} + 1 < \frac{n-1}{2}$. Thus there is an integer $\widetilde{k} \in (\sqrt{n} - 1,\ \sqrt{n} + 1)$ such that

$$\nu(\widetilde{k}) = \overline{V}_{\widetilde{k}} > \nu(\widehat{k}) \geq \overline{V}_{\widehat{k}}.$$

Since $\overline{V}_{\widetilde{k}} > \overline{V}_{\widehat{k}}$, we have obtained the desired contradiction and the proof of the theorem is complete.

□

Appendix B

In this appendix, we include the Mathematica code[13] that has been used to generate the numerical examples in this paper. The Mathematica notebook as it is reported here deals with the case $n = 9$ and $\delta = 0.6$, but any other values of these parameters can be chosen as well. The Mathematica notebook consists of three pages: On the first page, the relevant variables and equations are defined, and the parameter values for n and δ are set. On the second page, the relevant case distinctions are implemented. Finally, on the third page, the notebook creates a table in which the values of the variables, in particular of λ_k^+ and λ_k^- are reported for every $k = 1, \ldots, n-1$. For each of the values of k, it is checked whether λ_k^+ and λ_k^- are both non-positive. If this is true, an equilibrium has been found.

[13]The authors thank Tettje Halbertsma for her assistance in creating the Mathematica code for the numerical examples.

$$eqsV = \left\{V == \frac{k}{n-1} * X + \left(\frac{n-1-k}{n-1}\right) * \delta * W\right\};$$

$$eqsW = \left\{W == \frac{Y-V}{n-1}\right\};$$

$$eqsX = \left\{X == 1 - k * \delta * V - Max\left[0, \frac{\delta * (n-1-2k)}{2n}\right] * Y\right\};$$

$$eqsY = \left\{Y == \frac{k}{\delta * k + (1-\delta) * (n-1)}\right\};$$

$$eqs\lambda p1 = \left\{\lambda p == -\frac{k+1}{n-1} * \delta * V + \frac{X - \delta * W}{n-1}\right\};$$

$$eqs\lambda p2 = \left\{\lambda p == -\frac{k+1}{n-1} * \left(\delta * V - \delta * \frac{Y}{n}\right) + \frac{X - \delta * W}{n-1}\right\};$$

$$eqs\lambda m1 = \left\{\lambda m == \frac{k-1}{n-1} * \delta * V - \frac{X - \delta * W}{n-1}\right\};$$

$$eqs\lambda m2 = \left\{\lambda m == \frac{k-1}{n-1} * \left(\delta * V - \delta * \frac{Y}{n}\right) - \frac{X - \delta * W}{n-1}\right\};$$

```
variables = {V, W, X, Y, λp, λm};
```

```
n = 9;
δ = 0.6;
```

```
For[k = 1, k ≤ n - 1, k++,
 If[k == 1,
  {eqsλp = eqsλp2;
   eqsλm = {λm == 0};}];
 If[k ≥ 2 && k ≤ (n - 3)/2,
  {eqsλp = eqsλp2;
   eqsλm = eqsλm2;}];
 If[k == (n - 1)/2,
  {eqsλp = eqsλp1;
   eqsλm = eqsλm2}];
 If[k ≥ (n + 1)/2 && k < n - 1,
  {eqsλp = eqsλp1;
   eqsλm = eqsλm1;}];
 If[k ≥ n - 1,
  {eqsλp = {λp == 0};
   eqsλm = eqsλm1}];

 equations = Join[eqsV, eqsW, eqsX, eqsY, eqsλp, eqsλm];
 matrixformeqs = Normal[CoefficientArrays[equations, variables]];
 A = matrixformeqs[[2]];
 b = -matrixformeqs[[1]];
 sol[k] = LinearSolve[A, b];
 If[sol[k][[5]] ≤ 0 && sol[k][[6]] ≤ 0,
  equi[k] = "yes", equi[k] = "no"]; Print[eqsλp2]
]
```

$$\left\{\lambda p == \frac{1}{8}\,(-0.6\,W + X) + \frac{1}{4}\,(-0.6\,V + 0.0666667\,Y)\right\}$$

$$\left\{\lambda p == \frac{1}{8}\,(-0.6\,W + X) - \frac{3}{8}\,(0.6\,V - 0.0666667\,Y)\right\}$$

$$\left\{\lambda p == \frac{1}{8}\,(-0.6\,W + X) + \frac{1}{2}\,(-0.6\,V + 0.0666667\,Y)\right\}$$

$$\left\{\lambda p == \frac{1}{8}\,(-0.6\,W + X) - \frac{5}{8}\,(0.6\,V - 0.0666667\,Y)\right\}$$

$$\left\{\lambda p == \frac{1}{8}\,(-0.6\,W + X) - \frac{3}{4}\,(0.6\,V - 0.0666667\,Y)\right\}$$

$$\left\{\lambda p == \frac{1}{8}\,(-0.6\,W + X) - \frac{7}{8}\,(0.6\,V - 0.0666667\,Y)\right\}$$

$$\left\{\lambda p == -0.6\,V + \frac{1}{8}\,(-0.6\,W + X) + 0.0666667\,Y\right\}$$

$$\left\{\lambda p == \frac{1}{8}\,(-0.6\,W + X) - \frac{9}{8}\,(0.6\,V - 0.0666667\,Y)\right\}$$

```
resulttabel = {{"k", "Vk", "Wk", "Xk", "Yk", "λ+k", "λ-k", "equilibrium?"}};
For[k = 1, k ≤ n - 1, k++,
  resulttabel = Join[resulttabel, {Join[{k}, sol[k], {equi[k]}]}];
 ];

Print["n=", n, " ", "δ=", δ]
Print[resulttabel // TableForm]
```

n=9 δ=0.6

k	Vk	Wk	Xk	Yk	λ+k	λ-k	equilib:
1	0.118962	0.0180245	0.875991	0.263158	0.0946888	0.	no
2	0.192012	0.0328166	0.708979	0.454545	0.054322	-0.0755481	no
3	0.225408	0.046824	0.554265	0.6	0.0181488	-0.0419601	no
4	0.235435	0.0598563	0.434956	0.714286	-0.0384078	-0.0147646	yes
5	0.223098	0.0729192	0.330706	0.806452	-0.0645248	0.0310601	no
6	0.20613	0.0845279	0.257934	0.882353	-0.0823159	0.0513964	no
7	0.188684	0.0946577	0.207526	0.945946	-0.0943692	0.0660666	no
8	0.172414	0.103448	0.172414	1.	0.	0.0767241	no

```
ClearAll["Global`*"];
```

References

ANESI, V. (2010) Noncooperative Foundations of Stable Sets in Voting Games, *Games and Economic Behavior, 70,* 488–493.

ANSOLABEHERE, S., J.M. SNYDER, AND M.M. TING (2005), Legislative Bargaining Under Weighted Voting, *American Economic Review, 95,* 981–1004.

ANSOLABEHERE, S. J.M. SNYDER, A.B. STRAUSS, AND M.M. TING (2005), Voting Weights and Formateur Advantages in the Formation of Coalition Governments, *American Journal of Political Science, 49,* 550–563.

BANKS, J. AND J. DUGGAN (2000), A Bargaining Model of Collective Choice, *American Political Science Review, 94,* 73–88.

BANKS J. AND J. DUGGAN (2006), A General Bargaining Model of Legislative Policy–making, *Quarterly Journal of Political Science, 1,* 49–85.

BARON, D. AND J. FEREJOHN (1989), Bargaining in Legislatures, *American Political Science Review, 83,* 1181–1206.

BATTAGLINI, M. AND S. COATE (2007) Inefficiency in Legislative Policy-making, *American Economic Review, 97,* 118–149.

BATTAGLINI, M., S. NUNNARI, AND T. PALFREY (2014) Dynamic Free-riding with Irreversible Investments, *American Economic Review, 104,* 2858–71.

BOWEN, T.R. AND Z. ZAHRAN (2012), On Dynamic Compromise, *Games and Economic Behavior, 76,* 391–419.

BRITZ, V., P.J.J. HERINGS, AND A. PREDTETCHINSKI (2014), On the Convergence to the Nash Bargaining Solution for Action–dependent Protocols, *Games and Economic Behavior, 86,* 178–183.

DIERMEIER, D. AND P. FONG (2011), Legislative Bargaining with Reconsideration, *Quarterly Journal of Economics, 126,* 947–985.

DIERMEIER, D. AND P. FONG (2012), Characterization of the von Neumann–Morgenstern Stable Set in a Non–cooperative Model of Dynamic Policy–making with a Persistent Agenda–setter, *Games and Economic Behavior, 76,* 349–353.

FAHRENBERGER, T. AND H. GERSBACH (2007), Legislative Process with Open Rules, CER–ETH Working Paper 07/64, ETH Zurich.

DIERMEIER, D. AND T.J. FEDDERSEN (1998), Cohesion in Legislatures and the Vote of Confidence Procedure, *American Political Science Review, 92,* 611–621.

ERASLAN, H. (2002), Uniqueness of Stationary Equilibrium Payoffs in the Baron–Ferejohn Model, *Journal of Economic Theory, 148,* 2195–2222.

FALCONIERI, S. (2004), Lobbying and the Optimality of Delegation: Learning from Political Institutions, Working Paper.

FRÉCHETTE, G., J.H. KAGEL, AND S.F. LEHRER (2003), Bargaining in Legislatures: An Experimental Investigation of Open versus Closed Amendment Rules, *American Political Science Review, 97,* 221–232.

KALANDRAKIS, A. (2004), A Three-player Dynamic Majoritarian Bargaining Game, *Journal of Economic Theory, 116,* 294–322.

LARUELLE, A. AND F. VALENCIANO (2008), Noncooperative Foundations of Bargaining Power in Committees and the Shapley–Shubik Index, *Games and Economic Behavior, 63,* 341–353.

LEBLANC, W., J. SNYDER, AND M. TRIPATHI (2000), Majority–rule Bargaining and the Underprovision of Public Investment Goods, *Journal of Public Economics, 75,* 21–47.

MCCARTY, N. (2000), Proposal Rights, Veto Rights, and Political Bargaining, *American Journal of Political Science, 44,* 506–522.

OLESZEK, W. (2011), Congressional Procedures and the Policy Process, Fourth Edition (CQ Press, 1995).

PRIMO, D. (2003), Open vs. Closed Rules in Budget Legislation: A Result and an Application, Working Paper.

PRIMO, D. (2007), A Comment on Baron and Ferejohn (1989): The Open Rule Equilibrium and Coalition Formation, *Public Choice, 130,* 129–135.

RUBINSTEIN, A. (1982), Perfect Equilibrium in a Bargaining Model, *Econometrica, 50,* 97–110.

Chapter 6

Information Sharing in Democratic Mechanisms

Joint with Hans Gersbach
Accepted for publication in the *International Journal of Game Theory*

Abstract: We examine how democratic mechanisms can yield socially desirable outcomes in the presence of uncertainty about an underlying state of nature. In particular, we consider a class of decision–making procedures which we call "democratic mechanisms." We depart from a conventional mechanism design approach because we aim for democratic mechanisms to reflect some basic properties of decision–making in democracies. In particular, actual decisions are made by majority voting. The proposals to be voted upon are made by a selfish agenda–setter. Moreover, communication is limited to a binary message space (that is, voting Yes or No). We show how suitable democratic mechanisms can resolve uncertainty, reveal the state of nature, and implement the Condorcet winner. We demonstrate that this implementation result requires (at most) two voting stages regardless of the number of states or the number of alternatives. We also show that implementation requires a conditional privilege for a small representative subset of the population.

JEL classification: D62, D72, H40

Keywords: Democratic Mechanisms, Polling, Sampling, Public Goods, Voting, Information sharing

Acknowledgements: The research presented in this paper has benefited from valuable suggestions by two anonymous referees, discussions with Carlos Alos–Ferrer, Clive Bell, Patrick Bolton, Tilman Börgers, Faruk Gul, Jeanne Hagenbach, Martin Hellwig, César Martinelli, Stephen Morris, Klaus Nehring, Francesco Squintani, and from the comments of conference audiences at SING10 in Krakow, the 2015 EEA Meeting in Mannheim, the World Congress of Game Theory 2016 in Maastricht, the Social Choice and Welfare Conference in Lund 2016, and the CEPR Theory Workshop in Gerzensee 2016. Seminar participants *at universities in Davis, Heidelberg, Irvine, Munich, San Diego, and Zurich have provided valuable comments and suggestions on democratic mechanisms. A precursor of the current manuscript has appeared as CER–ETH Working Paper No. 14/199 by Britz and Gersbach (2014) under the title "Experimentation in Democratic Mechanisms."*

1 Introduction

Motivation and approach

The ability of democratic decision–making procedures to achieve socially optimal outcomes is at the core of a long–standing and complex debate with many unresolved issues. One challenging question, in particular, is how democratic mechanisms can help reveal information about underlying state variables.

There is an important literature that deals with *information aggregation* in a common value setting. That is, voters agree that there is a "best" alternative that should be selected. However, they are uncertain about which alternative is the best one. More specifically, each voter receives only a noisy signal of how good the alternatives are. For instance, members of a jury agree that a defendant should be convicted if he committed the crime, and acquitted otherwise. However, the different jurors may evaluate the available evidence differently. Likewise, members of a recruiting committee may all agree that they wish to hire the most suitable candidate for a job, but have different subjective impressions of who is the best candidate. In such settings, voting procedures can be a useful tool to aggregate the information dispersed among the voters. Ideally, this information aggregation fades out the noise in individual signals and thus reveals the underlying state variable. A classical result in this area is the *Condorcet Jury Theorem.* For seminal contributions to the literature on information aggregation, see Feddersen and Pesendorfer (1997) and Austen–Smith and Banks (1996).

Our paper differs from the information aggregation literature: In our model, citizens have private information about their own type. They are perfectly aware of their own preferences over the alternatives. Moreover, their type is also a noisy signal of the underlying state of nature, which determines the distribution of types in the population. In order to distinguish our work from the aforementioned literature on *information aggregation,* we use the term *information sharing* throughout. By sharing private information about their types, citizens can learn the state of nature and thus the type distribution. In the class of democratic mechanisms which we consider, knowledge about the type distribution is vital for the implementation of the Condorcet winner. We will see that some citizens may find it in their interest to hide or misrepresent information about their types, and thereby about the state.

We study situations in which individual benefits from public policies are privately observed realizations from a probability distribution which itself depends on an unknown state variable. In particular, citizens are uncertain not only of other citizens' types, but also of the underlying distribution of types in the population. To give one example, many societies face trade–offs between protecting citizens from terrorism and privacy concerns. Citizens differ in their preferences over these considerations. Each citizen may be uncertain about the extent to which other individual citizens weigh privacy vs.

counter–terrorism. Moreover, a citizen is also uncertain about the distribution of such preferences in the population.

In this paper, we consider a continuum society which faces the following collective choice problem: There is a finite number of feasible public good levels. An individual citizen's preferences over these public good levels depend on his privately observed type. An underlying state of nature determines the probability distribution from which each citizen's type is independently drawn. Each state of nature is associated with a Condorcet winner. That is, one feasible public good level is preferred by the majority to any other feasible public good level. Citizens have a common prior belief about the state of nature. Each citizen is privately informed about his own type, and therefore perfectly aware of his own preferences. If citizens share their private information, they learn the distribution of preferences in the population, and hence the true state.

We are interested in a class of collective choice procedures which we call "democratic mechanisms." They bear some resemblance to mechanisms with restricted message spaces as well as to voting games. More specifically, democratic mechanisms share information in a similar way as studied in the mechanism design literature. However, decisions in democratic mechanisms are made by majority voting. We intend democratic mechanisms to reflect some important principles and features of real democratic processes:

- Voters are typically not able to communicate their preferences in full detail. Instead, they may only be able to vote in favor or against a proposal, or choose between, say, a Republican and a Democrat. The notion of a democratic mechanism reflects this by allowing only a binary message space.

- Arguably, the most basic principle of democratic decision–making is that the implementation of an alternative requires approval by a majority. Indeed, our model of democratic mechanisms assumes that a status quo can only be replaced by a new policy if the majority of citizens agree. This is in contrast to a standard mechanism design approach, where citizens only reveal information, but do not explicitly make the actual decision.

- Actual democratic processes are often shaped by elites (politicians, lobbyists, activists) who are selfish and use their power over the political process for their own benefit. Our notion of a democratic mechanism reflects this by letting a selfish "agenda–setter" choose the feasible alternative which is then voted upon.

Binary messages, majority voting, and a selfish agenda–setter are the three features that define our notion of a democratic mechanism. We aim for democratic mechanisms which involve as few voting stages as possible, and are therefore "procedurally efficient." The rationale is that, in practice, it is costly to organize an election or a referendum.

In the present paper, we ask which democratic mechanisms reliably reveal the state of nature and implement the Condorcet winner in the aforementioned collective choice problem. In order to accomplish implementation, a democratic mechanism must overcome two obstacles, which we refer to as *manipulation* and *exploitation.* We discuss these two obstacles in turn:

First, in order to discover the Condorcet winner, citizens need to share their private information using binary messages. A standard problem in mechanism design is the presence of incentives to strategically misrepresent one's private information, thus *manipulating* the mechanism. Not surprisingly, this obstacle to information sharing is also present in democratic mechanisms. In addition, due to the agenda–setter's active role, there is a second obstacle to the desired implementation result: Once the state of nature has been revealed, the agenda–setter may want to use this knowledge for his own benefit, rather than work "in good faith" towards the implementation of the Condorcet winner. Such strategic behavior by the agenda–setter is called *exploitation* throughout this paper.

Main results

We show an impossibility result and an existence result: First, we focus on the simplest possible ("baseline") democratic mechanisms, which rely on a single voting stage preceded by binary communication. We show that a baseline democratic mechanism fails to generally implement the Condorcet winner. We distinguish between the case where implementation fails because citizens *manipulate* information sharing, and the case where implementation fails because the agenda–setter *exploits* information sharing. Second, we derive an existence result, that is, we construct a democratic mechanism which is immune to both the voters' and the agenda–setter's attempts to manipulate or exploit information sharing, and which therefore implements the Condorcet winner. This democratic mechanism relies on two features: It grants a conditional privilege to a small random sample of the population, and requires a two–stage voting procedure. These two features are essential for implementation.

At a more detailed level, the mechanism we develop can be described as follows: A small representative sample of the population is drawn. Members of the sample group coordinate their binary messages so as to communicate information about the state. An agenda–setter observes the sample group's messages and proposes a public good level. This proposal is decided upon in two voting stages: First, the entire population makes a choice between the proposal and the highest feasible alternative which is lower than the proposal. Second, the winner of this vote is pitted against the status quo. Sample group members as well as the agenda–setter are exempted from taxation if and only if the proposal made by the agenda–setter prevails in both voting stages.

We show that this democratic mechanism guarantees the implementation of the Condorcet winner by eliminating all bad incentives for both *exploitation* and *manipulation.* Our democratic mechanism yields truthful information sharing, optimal public good provision, and consequently the implementation of the Condorcet winner in two voting stages, regardless of the number of states or alternatives. Our democratic mechanism requires granting a conditional privilege to a small subset of citizens. This privilege involves a tax–exemption as well as, possibly, a transfer. Due to this conditional privilege, the democratic mechanism we propose may lead to an allocation of utilities that is close, but not exactly equal, to utilities that are stable to majority voting. In a nutshell, in order to find and implement the Condorcet winning policy *alternative,* our democratic mechanism must allow for a small distortion away from the Condorcet winning *allocation.* This distortion, however, is arbitrarily small.

Contribution and literature

Our contribution lies at the interface of mechanism design, social choice, and constitutional choice. As in mechanism design theory, we consider an environment with privately informed agents, and our aim is to find game forms which accomplish a socially desirable collective choice through truthful revelation of private information. Moreover, our focus on democratic mechanisms is related to social choice theory in two ways: First, it makes decisions by voting. Second, our approach pays attention to both the voters' and the agenda–setter's incentives. Finally, we evaluate the democratic mechanism devised in our paper against the requirement that citizens be treated equally, which is central to constitutional choice.

A closely related recent contribution is Bierbrauer and Hellwig (2016). Their work is a starting point for our research in two respects: First, we follow Bierbrauer and Hellwig in considering decision–making on public good levels through voting procedures. Second, we allow groups of citizens to coordinate their signals, and establish an implementation result which is robust to such coordinated behavior. This uses one of Bierbrauer and Hellwig's main insights, namely that "coalition–proofness" is a desirable property of an incentive mechanism for public good provision. However, we focus on different issues than Bierbrauer and Hellwig. In particular, we examine whether and how the Condorcet winner can be implemented by voting procedures in the presence of two complications: First, in contrast to the mechanism design approach, proposals are made by a selfish agenda–setter and there is no mechanism designer. Second, our model includes uncertainty about the type distribution. This uncertainty has to be resolved in order to proceed with implementation.

Our objective is to implement the Condorcet winner, which need not coincide with the social welfare optimum. The rationale is that we are interested in what can be accomplished by a class of *democratic* procedures. It seems innocuous that a democratic system should select an alternative that the majority prefers to any other alternative. The Condorcet winner is the only alternative that is renegotiation–proof. Nevertheless, this requirement is non–trivial: Many electoral systems do not satisfy it. For example, in elections with a run–off round, the Condorcet winner may fail to be selected. Therefore, it is interesting to study democratic mechanisms which do reliably select the Condorcet winner.

One could have in mind richer mechanisms in which the required (super–) majority depends on the alternative that is voted upon, as in the recent work of Bierbrauer and Hellwig (2016). With such more complex rules, one could aim for democratic mechanisms which find the social welfare optimum rather than the Condorcet winner.

Moreover, we could distiniguish between critical and day-to-day decisions by either introducing fixed costs for any positive level of public good provision or allowing for different types of public good provision. Critical decisions would then be about whether to provide the public good at all or which type of public good to provide. For such cases, Aidt and Giovannoni (2011) have outlined a theory how different collective decision rules can be used for different types of decisions. This theory could be applied in our context as well.

In this paper, we are interested in the implementation of the Condorcet winner through a class of collective decision procedures called democratic mechanisms. Of course, there may be alternative ways how one could accomplish this goal. For instance, a social planner could identify the Condorcet winner by eliciting citizens' pairwise preferences over all the alternatives. The procedure we propose has two advantages over the elicitation of all pairwise preferences: First, we want to let go of the idea that there is a benevolent social planner who designs a scheme to elicit citizens' preferences. Instead, we allow for a selfish agenda–setter. Second, we are interested in procedures that are efficient in the sense that they require as few voting rounds as possible. Our contribution is therefore not that we show how one can arrive at the Condorcet winner. Rather, we have studied how the Condorcet winner can be implemented using a class of collective decision procedures that allows for the agenda to be shaped by selfish members of society and that is procedurally efficient. Setting the right incentives for the selfish agenda–setter and curbing his power to exploit the mechanism for his own self–interest is a crucial challenge in our paper: Depending on their types, some agenda–setters may want to propose too little or too much public good provision. The presence of a selfish agenda–setter and the desired procedural efficiency complicate the problem at hand despite the fact that the Median Voter Theorem holds.

Our work fits into the context of a vast body of literature on optimal constitutions which has developed after the classic work of Buchanan and Tullock (1962). Aghion and Bolton (2003) have introduced incomplete social contracts and have explored how simple or qualified majority rules balance the need to overcome vested interests and respect majority preferences. Gersbach (2009) has shown how increasingly sophisticated combinations of agenda rules, treatment rules, and decision rules can yield first–best allocations when each citizen faces only two possible realizations: being either a winner or a loser of a public project. The present paper allows for uncertainty about the distribution of valuations. Neither individual valuations nor the underlying distribution are common knowledge. In particular, we explore the scope of democratic mechanisms with minimal message spaces but with rich type spaces and uncertainty about type distribution.

In the literature on collective decision–making following the famous Gibbard–Satterthwaite theorem, one central issue has been manipulation, which can occur in various forms. On the one hand, an extensive literature has focused on manipulation through strategic voting. For some recent contributions in this realm, see, for instance, Bouton (2013), Cho (2014), Dellis (2013), Gershkov et al. (2017), Rasch (2014), Van der Straeten et al. (2010).) On the other hand, manipulation of collective decision–making can also occur through communication, as the literature on polling has demonstrated. Meirowitz (2005) analyzes the case of two candidates competing for political office, who attempt to elicit voter preferences through polling. Morgan and Stocken (2008) discuss to what extent policy–makers can learn from polling. A seminal contribution to the literature on strategic communication is Crawford and Sobel (1982). The limitations of polling are also discussed by Goeree and Grosser (2007) and Taylor and Yildirim (2005). Bernhardt, Duggan, and Squintani (2008) provide a recent survey of the polling literature. One contribution of the present paper is that it studies the implementation of the Condorcet-winning alternative in a model where *both* strategic voting *and* strategic communication are possible.[1]

The paper is organized as follows. In Section 2, we formally introduce the collective choice problem faced by our model society. In Section 3, we formalize the notion of a democratic mechanism. In Section 4, we consider the simplest possible ("baseline") democratic mechanism, and demonstrate how it is prone to exploitation and manipulation. Then, we proceed to introduce a democratic mechanism which reliably accomplishes information sharing and guarantees the implementation of the Condorcet winner. We discuss the main implementation result in Section 5. In Section 6, we discuss a simplified implementation result which holds on a subclass of public good problems. Section 7 concludes.

2 The collective choice problem

We consider a collective choice problem in a society which consists of a continuum of risk–neutral citizens of unit mass. Citizens can choose how much of a public good should be provided. However, our main results and conclusions would translate into a more general setting with the following characteristics: Feasible alternatives are discrete and can be

[1]Another relevant branch of literature is that on policy experimentation. It refers to situations in which new policies are implemented and "tested" in one constituency so that the entire society can learn from experience. After the seminal contribution of Rose–Ackermann (1980), this line of research has been extended and deepened by Kollman, Miller, and Page (2000), Strumpf (2002), Volden (2006), Shipan and Volden (2006), Volden, Ting, and Carpenter (2008), Cai and Treisman (2009), Bednar (2011), as well as Callander and Harstad (2013). Our work differs from this literature in that we only allow policy decisions to be made for the entire society at once.

ordered along a single dimension, and preferences are single–peaked.[2]

Each citizen is privately informed about his type z. The type space Z is a closed, non–empty, and non–degenerate interval in $\mathbb{R}_{++}$.[3] We refer to a citizen of type z as *citizen* z. We denote the interior of the type space by $int(Z)$. The provision of a public good level $q \in \mathbb{R}_+$ comes at a cost $c(q)$ to each citizen. The cost function $c(q)$ is strictly increasing and strictly convex, and $c(0) = 0$. A citizen's benefit from a unit of public good provision is given by his type. More specifically, we can write citizen z's utility from public good provision as $u(z, q) = zq - c(q)$.

The model features both individual and collective uncertainty. We assume that there are finitely many states of nature; the state space is $N = \{1, \dots, n\}$. We will often use i or k to index the elements of N. If the state of nature is k, then citizens' types are independent draws from a probability distribution on Z with density f_k and cumulative distribution function F_k. Citizens share a common prior belief p about the state of nature, where $p_k > 0$ for every $k \in N$. Due to Bayesian updating, citizen z believes in state k with probability $\beta_k(z) = \frac{f_k(z)p_k}{\sum_{j=1}^{n} f_j(z)p_j} > 0$.

We make the following assumptions:

Assumption 1.

1. *For any $z \in int(Z)$, we have $F_1(z) > \ldots > F_n(z)$.*
2. *For every $k \in N$, there is a Condorcet winner, that is, there is $q_k > 0$ such that a majority of citizens prefers q_k to any $q \in \mathbb{R}_+ \setminus \{q_k\}$.*
   Moreover, it holds that $0 < q_1 < \ldots < q_n$.
3. *For every $k \in N$ there is some $\widetilde{q}_k > 0$ such that a majority of citizens strictly prefers a public good level $q > 0$ to zero public good if and only if $q < \widetilde{q}_k$, and a majority strictly prefers zero public good to $q > 0$ if and only if $q > \widetilde{q}_k$.*
4. *For every $k \in N$ and every $z \in Z$, it holds that $\beta_k(z) > 0$ and Bayesian updating is monotone.*[4]

In Appendix A, we explain how Assumption 1 can be deduced from a set of deeper model assumptions on the cost function, type space, and probability distributions.

In the remainder of the paper, we restrict attention to a discrete set Q of feasible public good levels. We assume that $\{0\} \cup \{q_1, \dots, q_n\} \subset Q$. We consider zero public good provision as a status quo, and we require the Condorcet winners to be feasible. Moreover, it is convenient to assume $(0, q_1) \cap Q = \emptyset$.

[2]In our formal model, we assume that the status quo is zero public good provision. In Subsection 5.3, we briefly discuss the possibility of a more general status quo.

[3]Our main results would carry over to the case where Z is only restricted to lie in $\mathbb{R}_+$. However, this would add a number of technical complications.

[4]Monotone Bayesian updating means that for any $z_1, z_2 \in Z$ with $z_1 < z_2$, the posterior probability distribution $\{\beta_k(z_2)\}_{k=1}^{n}$ stochastically dominates $\{\beta_k(z_1)\}_{k=1}^{n}$.

For any $q' \in (0, q_n)$, let

$$\begin{aligned} \psi_-(q') &= \max\{q \in Q | q < q'\}, \\ \psi_+(q') &= \min\{q \in Q | q > q'\}. \end{aligned}$$

To sum up, the state space N, the type space Z, the family of cumulative distribution functions F, the feasible set Q, the cost function c, and the common prior p constitute a choice problem, which we denote by P. The set of all such problems is denoted by $\mathcal{P}$. For each problem $P \in \mathcal{P}$, it is commonly known that a Condorcet winner exists. However, no individual citizen knows which alternative is the Condorcet winner. In what follows, we will be interested in decision–making procedures which reliably implement the Condorcet winner for the whole class $\mathcal{P}$.

It is important to note that suitably coordinated binary messages suffice to share information and reveal the Condorcet winner. To be more specific, we can interpret $F_k(z)$ as the cross–sectional distribution of z in the population when the state is k. It is well–known that this interpretation requires the application of a suitable version of the law of large numbers for a continuum of random variables. More specifically, for any public good problem $P \in \mathcal{P}$ with type space Z, state space N, cumulative distribution functions F, and set of feasible alternatives Q, fix some $z^P \in int(Z)$, and assume that all citizens $z \geq z^P$ send one message (say "Yes") and all citizens of type $z < z^P$ send another message (say "No"). Denote by δ the observed measure of citizens sending message "Yes." Then, define $\varphi^P : [0,1] \to Q$ as follows:

$$\varphi^P(\delta) = \begin{cases} q_1 & \text{if } 0 \leq \delta < 1 - F_2(z^P), \\ q_k & \text{if } 1 - F_k(z^P) \leq \delta < 1 - F_{k+1}(z^P), \text{ for } k = 2, \ldots, n-1, \\ q_n & \text{if } 1 - F_n(z^P) \leq \delta \leq 1. \end{cases}$$

Recall that we have assumed that $F_1(z) > \ldots > F_n(z)$ for any $z \in int(Z)$. Thus, learning the share of citizens whose type is higher than some critical (interior) type suffices to learn the state. More precisely, if all citizens $z \geq z^P$ send one message, and all citizens $z < z^P$ send the other message, then the map φ^P reveals the Condorcet winner. The critical type z^P can be chosen arbitrarily for each public good problem.

The question arises how one could accomplish coordination on a particular value of z^P. In the sequel of the paper, we will discuss democratic mechanisms in which an agenda–setter has a rich message space, while other citizens can only communicate by sending binary messages. It is therefore possible for the agenda–setter to establish coordination by simply asking, "Is your type higher or lower than z^P?" for some z^P of his choice.

3 Democratic mechanisms

In the literature, a *mechanism* is commonly defined as a map from messages to outcomes. The decision–making process is as follows: First, a designer announces the mechanism. Then, citizens report their types. Finally, the outcome associated to the reported type profile by the mechanism is implemented. This mechanism design approach hinges on several tacit assumptions: First, it requires the message space to be as rich as the type space. Second, it assumes that the designer is benevolent with regard to some social objective, and that he is committed to the mechanism he has designed. In particular, after observing the messages, the designer cannot "change his mind" and choose an alternative different from the one prescribed by the mechanism. Third, the mechanism design approach assumes a kind of coercive power of the designer: The outcome of the mechanism can be implemented without the citizens' explicit consent.

Our argument is that there is a discrepancy between these tacit assumptions in the mechanism design approach and observed features of the democratic process. It is our aim to bridge this gap. In particular, we modify the standard notion of a mechanism in three ways:

- First, we allow only binary messages. This reflects the idea that, in a democratic decision–making process, citizens are often unable to express their preferences in full detail.
- Second, we allow a selfish agenda–setter to intervene in the decision–making process. In democratic systems, the process is shaped by politicians, activists, lobbyists, or civil servants who have preferences of their own.
- Third, we require the explicit consent of a majority of citizens before the status quo can be replaced by any alternative outcome. This is arguably the most basic condition one would want to impose on a democratic process.

There is a natural tension between the agenda–setter's selfishness and the requirement of majority voting. In particular, the need for approval by the majority could be seen as balancing the power of the selfish agenda–setter.

We are now going to give a more formal description of the kind of decision–making process that we refer to as a *democratic mechanism.* It can be decomposed into two parts: First, private information is communicated and shared through binary messages. Second, a decision is taken through voting. More specifically, decision–making in a democratic mechanism proceeds as follows:

Citizens simultaneously send binary messages. We denote the binary message space by $\{Yes, No\}$. We will often refer to the message "Yes" as the *positive signal,* and to the message "No" as the *negative signal.* We denote the observed share of citizens who have sent the positive signal by $\delta \in [0,1]$. For every public good problem $P \in \mathcal{P}$, an *information mapping* φ^P designates a feasible public good level given the observed share of positive signals. In mapping reports to outcomes, the information mapping bears a resemblance to a "mechanism" in the conventional sense (albeit with a restricted message space). The crucial difference is that the communication entering the information mapping is mere cheap talk. The alternative designated by φ^P is not "automatically" implemented. Instead, after messages have been sent and δ has been observed, there is a voting procedure V in which a selfish agenda–setter makes a proposal that may or may not coincide with $\varphi^P(\delta)$. The status quo can only be replaced by an alternative if the majority explicitly approves it.[5] However, the voting procedure V may specify any number of additional voting stages before this final decision is made. Let $\varphi = (\varphi^P)_{P\in\mathcal{P}}$. Then, the pair (φ, V) constitutes a *democratic mechanism*.[6]

Like a "mechanism" as conventionally understood, a democratic mechanism involves a mapping from messages to outcomes. In addition, however, it involves voting on the suggested outcome as well as agenda–setting by a selfish agent.

Given a democratic mechanism (φ, V) and a public good problem $P \in \mathcal{P}$ (with type space Z, feasible set $Q, \dots$), let a *communication strategy* $\sigma^P : Z \to \{Yes, No\}$ assign a message to each type. Similarly, let a *proposal strategy* $\rho^P : Z \times [0,1] \to Q$ assign a proposal to the agenda–setter's type and to the observed share of positive signals.

We say that citizens can *manipulate* information sharing in the democratic mechanism (φ, V) if there is $P \in \mathcal{P}$ such that the following holds: For every pair[7] (ρ^P, σ^P) such that

$$\varphi^P(\delta_i(\sigma^P)) = \rho^P(z, \delta_i(\sigma^P)) = q_i, \ \forall i \in N, \ \forall z \in Z,$$

there is a subset $\widehat{Z} \subset Z$ and a communication strategy $\widehat{\sigma}^P$ such that (i) $\widehat{\sigma}^P(z) = \sigma^P(z)$ for all $z \in Z \setminus \widehat{Z}$, while $\widehat{\sigma}^P(z) \neq \sigma^P(z)$ for some $z \in \widehat{Z}$; and (ii) the strategy profile $(\rho^P, \widehat{\sigma}^P)$ makes all members of $\widehat{Z}$ weakly better off[8], and makes some members of $\widehat{Z}$ strictly better off than the strategy profile (ρ^P, σ^P). The problem of manipulation is familiar from mechanism design theory. We note that the above definition of manipulation allows

[5]For our results, it is not crucial what is assumed to happen if exactly half of the citizens vote for either option. We can assume throughout the paper that ties are broken by fair randomization.

[6]We stress that a democratic mechanism (φ, V) is defined independently of a particular $P \in \mathcal{P}$. The reason is that we aim to implement the Condorcet winner on the whole set $\mathcal{P}$.

[7]Due to the construction of φ^P, the communication strategy σ^P which ensures $\varphi^P(\delta_i(\sigma^P)) = q_i$ for every $i \in N$ has citizen z send the positive signal if and only if $z \geq z^P$. However, there are several proposal strategies ρ^P such that $\rho^P(z, \delta_i(\sigma^P)) = q_i$ for every $i \in N$. In particular, these proposal strategies can prescribe different proposals for $\delta \in [0,1] \setminus \{\delta_1(\sigma^P), \dots, \delta_n(\sigma^P)\}$.

[8]Whether sending different messages can make citizens "better off" clearly depends on the voting decisions that citizens expect to be made when the democratic mechanism proceeds to the voting procedure V. In accordance with the idea of backward induction, we assume here that citizens anticipate voting decisions when they choose their messages. Likewise, the agenda–setter anticipates the voting decisions when making a proposal. For our analysis, only "sincere voting" decisions (in a sense to be made precise) will be relevant.

citizens to participate in a coordinated deviation even if they have no *strict* incentive to do so. This definition of manipulation makes it harder to discover the state and implement the Condorcet winner, and thus makes our existence result stronger. A narrower definition of manipulation (that is, requiring that citizens only participate if they become strictly better off), however, would not undo the impossibility result stated in Proposition 1. Moreover, the above definition of manipulation is also quite broad in the sense that it allows for coordination between any subset of the citizens, as long as each member becomes weakly better off.

In order to derive the impossibility result, we consider the case where a group of citizens sharing the same preference ranking over alternatives coordinate their actions in order to manipulate information sharing. Since there is a discrete set of alternatives, citizens sharing the same preference ranking need not all have the exact same marginal utilities. We note that the above definition of manipulation would even allow coordination within larger groups of citizens.

We say that the agenda–setter can *exploit* information sharing in the democratic mechanism (φ, V) if there is $P \in \mathcal{P}$ such that the following holds: For every pair (ρ^P, σ^P) of a proposal and a communication strategy such that

$$\varphi^P(\delta_i(\sigma^P)) = \rho^P(z, \delta_i(\sigma^P)) = q_i, \ \forall i \in N, \ \forall z \in Z,$$

there are a state $i \in N$ and a type $z \in Z$ such that, upon observing $\varphi^P(\delta_i(\sigma^P))$, citizen z is better off proposing some $\widehat{q} \in Q \setminus \{q_i\}$ rather than proposing q_i. The problem of exploitation is not present in standard mechanism design theory.

We say that a democratic mechanism (φ, V) *implements* the Condorcet winner if the following holds: For every public good problem $P \in \mathcal{P}$, there are strategies (ρ^P, σ^P) such that (i) $\varphi^P(\delta_i(\sigma^P)) = \rho^P(z, \delta_i(\sigma^P)) = q_i$ for every $i \in N$ and every $z \in Z$, (ii) citizens cannot manipulate information sharing, and (iii) the agenda–setter cannot exploit information sharing.[9]

4 Baseline democratic mechanism

The simplest possible voting procedure is the following: After observing the messages sent by citizens, a selfish agenda–setter makes a proposal. Citizens decide by simple majority voting whether to implement this proposal or stick to the status quo. A democratic mechanism which involves this voting procedure (and some information mapping) is called a *baseline democratic mechanism.* In a baseline democratic mechanism, the decision to vote in favor of the proposal or the status quo is binary and final. Hence, there is no room

[9]Our definition of implementation requires that the Condorcet winner is implemented regardless of the agenda–setter's type. One implication is that our implementation result would hold in a setup where the agenda–setter is randomly chosen from the entire population.

for strategic behavior during the voting procedure. Indeed, *sincere voting* is optimal in a baseline democratic mechanism. That is, if the agenda–setter has made the proposal $q \in Q$, then citizen z votes in favor of q if and only if he prefers it to zero public good provision, which is the status quo. More formally, citizen z votes in favor of q if and only if $z \geq c(q)/q$.

We want to show that baseline democratic mechanisms do not generally implement the Condorcet winner. In the previous section, we have already hinted at the two conditions for implementation of the Condorcet winner: The citizens should have no incentive to *manipulate* information sharing, and the agenda–setter should have no incentive to *exploit* information sharing, whatever his type is. For baseline democratic mechanisms, we establish impossibility in two ways: First, Proposition 6.1 below shows that whenever information sharing is not manipulated, it can be exploited. Then, we take the opposite perspective: Suppose that the agenda–setter does not exploit information sharing. Proposition 6.2 claims that, under an additional condition on the public good problem, a subset of citizens can successfully manipulate information sharing if the agenda–setter does not exploit it.

Proposition 6.1. *In a baseline democratic mechanism, if citizens do not manipulate information sharing, then the agenda–setter can exploit it. In particular, exploitation is possible for every $P \in \mathcal{P}$.*

The proof of Proposition 6.1 can be found in Appendix B.

Our definition of implementation requires that exploitation is impossible for every $P \in \mathcal{P}$. Proposition 6.1 not only says that there is $P \in \mathcal{P}$ for which exploitation is possible. Actually, we have shown that, even in the absence of manipulation, exploitation can prevent implementation in *every* public good problem $P \in \mathcal{P}$.

Now we turn to manipulation of information sharing by citizens. We allow groups of citizens with the same preference ranking to coordinate their signals. We allow manipulations that depend on cooperation of citizens who only weakly benefit. This is a conservative assumption that biases our results against implementation.

Given that we work with a continuum society, it is not meaningful to consider "unilateral deviations" by individual citizens. Instead, we have assumed that citizens who have the same preference ranking over the feasible alternatives can coordinate their actions. For the implementation of the Condorcet winner, we require that it is robust to deviations by such a group of citizens. This approach is in line with one of the main insights of a recent paper by Bierbrauer and Hellwig (2016), who argue that such "coalition–proofness" is a desirable property of an incentive mechanism. The assumption that groups of citizens with the same preference ranking can coordinate their moves is clearly more conservative than allowing only deviations by a smaller group or by individuals. One might wonder whether groups of citizens who do not all have the same preference ranking could also coordinate their actions.

In particular, we will consider a subset of citizens denoted by Z_-. This set $Z_- \subset Z$ contains all types lower than $c(q_1)/q_1$. Those are the citizens who consistently prefer lower public good levels to higher ones. Continue to suppose that information is shared by having citizen z send the positive signal if and only if $z \geq z^P$, for some $z^P > c(q_1)/q_1$. Citizens in Z_- may be able to manipulate this information sharing in the following way: Some share of members of Z_- could send positive instead of negative signals. In this way, they could mimic the vote share associated with a higher state than the actual state. Thereby, they would prompt the agenda–setter to make a proposal which is "too high." This proposal would then lose against the status quo, resulting in zero public good provision.

To be more precise, let us consider public good problems with the following properties:

Definition 6.1.

1. *The public good problem $P \in \mathcal{P}$ has the* distance property *if $\widetilde{q}_i < q_{i+1}$ for every $i = 1, \ldots, n-1$.*

2. *Suppose that the public good problem $P \in \mathcal{P}$ has the distance property. Then, a state $i \in N$ is* concealable *if $F_i(z) - F_{i+1}(z)$ is a quasi–concave function of z, and attains its unique maximum at a point in Z_-.*

The intuition behind the distance property is as follows: Due to the single–peaked preferences, if the true state is i, and the agenda–setter "overshoots" by proposing a quantity somewhat greater than q_i, then a majority would prefer that greater quantity to the status quo. If the scope for "overshooting" is so great that even q_{i+1} would be preferred by the majority over the status quo, then the distance property is violated.

Let us briefly discuss possible economic interpretations of the "distance property." For instance, consider the example of a school district that decides on the number of teachers it hires. This problem would not have the distance property: The increments between feasible alternatives are (nearly) arbitrarily small. After all, there is no constraint which says that teachers must be hired in units of, say, a hundred teachers. As a counter–example, consider the case of building a new bypass road around a city. In such a case, the available options could be (i) building no new road at all, (ii) building the road as a single carriageway, or (iii) building the road as a dual carriageway. This problem would have the distance property: The costs and benefits differ a lot between the different options, and there is no easy way to introduce additional options which differ only incrementally from the available alternatives. After all, building only a short piece of a bypass road is not useful, nor are arbitrarily small stretches of dual carriageways. Similar examples are quite typical of public infrastructure projects.

The next proposition claims that, in a baseline democratic mechanism, even if the agenda–setter does not exploit information sharing, implementation fails due to manipulation when the distance property holds and a state is concealable.

Proposition 6.2. *In a baseline democratic mechanism, if the agenda–setter does not exploit information sharing, then citizens can manipulate it. In particular, manipulation is possible for every public good problem $P \in \mathcal{P}$ which satisfies the distance property, and in which at least one state is concealable.*

The proof of Proposition 6.2 can be found in Appendix B. We now provide a numerical example which illustrates the distance property and the concealable state.

Example 6.1. *Fix some $\varepsilon > 0$. Let the type space be $Z = [\varepsilon, \varepsilon + 1]$, and the cumulative distribution functions*

$$\begin{aligned} F_1(z) &= z - \varepsilon, \\ F_2(z) &= (z - \varepsilon)^{1.2}. \end{aligned}$$

Moreover, let the cost function be $c(q) = q^{1.1} + \varepsilon q$. This implies that marginal cost is $c'(q) = 1.1(q^{0.1}) + \varepsilon$ and average cost $c(q)/q = q^{0.1} + \varepsilon$.

Using the equality $F_k(c'(q_k)) = 1/2$, we can compute the Condorcet winners

$$\begin{aligned} q_1 &= \left(\frac{0.5}{1.1}\right)^{10} \approx 0.000377, \\ q_2 &= \left(\frac{0.5^{(1/1.2)}}{1.1}\right)^{10} \approx 0.0012. \end{aligned}$$

Moreover, we can use the equality $F_1(c(\widetilde{q}_1)/\widetilde{q}_1) = 1/2$ to compute

$$\widetilde{q}_1 = (0.5)^{10} \approx 0.000977.$$

Indeed, we observe that $q_1 < \widetilde{q}_1 < q_2$, that is, the distance property holds in this example.

It is easy to check that the vertical distance $F_1(z) - F_2(z)$ is maximized at the point

$$z^* = \left(\frac{1}{1.2}\right)^5 + \varepsilon \approx 0.402 + \varepsilon.$$

The set Z_- contains all types lower than

$$c(q_1)/q_1 = (q_1)^{0.1} + \varepsilon = \frac{0.5}{1.1} + \varepsilon \approx 0.4545 + \varepsilon.$$

We observe that the inequality $z^* < c(q_1)/q_1$ holds, that is, state 1 is concealable in this example.

The analysis and propositions in this section amount to an impossibility result: The baseline democratic mechanism does not generally implement the Condorcet winner. We note that either Proposition 6.1 or Proposition 6.2 alone imply this impossibility. Hence, it is not the interaction of exploitation and manipulation which leads to the failure of implementation. Even if either exploitation or manipulation could be eliminated, implementation of the Condorcet winner would remain impossible with a baseline democratic mechanism.

In the next section, we extend democratic mechanisms beyond the "baseline" in such a way that implementation of the Condorcet winner is accomplished.

5 Implementation result

In this section, we show how the baseline democratic mechanism can be modified in order to obtain the desired implementation of the Condorcet winner. In particular, two modifications of the baseline democratic mechanism are needed: First, voting proceeds in two stages rather than just one stage. Second, a small representative sample of the population (as well as the agenda–setter) is granted a conditional privilege.

The subsequent derivations also reveal that both modifications are necessary to obtain the general implementation result on our theorem. Privileges for the sample group and the agenda–setter are needed to ensure that they can only exploit or manipulate information sharing by aiming for a public good level higher than the Condorcet winner. The two–stage voting procedure is a corrective by which citizens outside the sample group can prevent excessive public good provision. Therefore, the combination of the (conditional) privileges and the two–stage voting procedure is essential for the implementation result. Neither a scheme of privileges nor a voting procedure with several stages alone can overturn the impossibility result described in the previous section.

5.1 Description of the mechanism

We now give the formal description of the *democratic mechanism with sampling and two–stage voting.*

Information mapping. A sample group of size $\lambda \in (0, 1)$ is randomly drawn from the population. Sample group members simultaneously send binary messages. The share δ of sample group members who have sent the positive signal is observed. When the public good problem is $P \in \mathcal{P}$, one feasible alternative is designated by the map φ^P.

Voting procedure. An agenda–setter determines the proposal $q \in Q$. The voting procedure has two stages: First, there is a selection stage in which citizens choose either the proposal q or the amendment $\psi_{-}(q)$. If the share of citizens who choose q is at least $(1 + \lambda)/2$, then q is selected. Otherwise, $\psi_{-}(q)$ is selected. Second, there is a voting stage in which citizens choose either the selected alternative from the previous stage, or the status quo. In order to win, the selected alternative requires a super–majority of $(1+\lambda)/2$ if the selected alternative is q, or a simple majority if the selected alternative is $\psi_{-}(q)$. [10]

Conditional privilege. If and only if the proposal q is finally implemented, then the agenda–setter and the sample group members are tax–exempt and, in addition, receive a transfer of τ.

The intuition behind this democratic mechanism is as follows: On the one hand, the privileges granted to the agenda–setter and sample group members give them an incentive to aim for a high public good level, regardless of their type. On the other hand, the two–stage voting procedure and the conditional character of the privilege allows citizens to rebuke any attempt by the "privileged few" to implement an excessive public good level.

The following theorem is the main implementation result of this paper.

Theorem 6.1. *With λ close to zero, the democratic mechanism with sampling and two–stage voting implements the Condorcet winner if τ is sufficiently large.*

The proof of Theorem 6.1 is relegated to Appendix C.

Since the sample contains the same information as the entire population, this eliminates any incentives by members of the sample to manipulate information sharing. However, it does give rise to an incentive for the agenda–setter to exploit information sharing. In order to counteract this problem, the tax–exemption is conditioned on the outcome of the selection stage. At that stage, the entire population is given an opportunity to make a minimal downward correction of the proposal. Given that τ is sufficiently high, this eliminates the incentive to exploit information sharing.

For a better understanding of Theorem 6.1 and its proof, let us briefly discuss what it means for the transfer τ to be "sufficiently large." In the proof of Theorem 6.1, we have had to check whether it can be profitable for the agenda–setter or sample group members to prevent the implementation of the Condorcet winner, say q_i. If they had such a profitable deviation, it would result in some quantity $\widehat{q}_i \neq q_i$ being implemented instead. In the proof, we have shown that this is only possible if $\widehat{q}_i$ has been an amendment at the selection stage. Hence, the agenda–setter and sample group would not receive any privileges under their deviation, and so an agenda–setter or sample group member of type z would receive a payoff of $z\widehat{q}_i - c(\widehat{q}_i)$. Without the deviation, the Condorcet winner would be implemented, and agenda–setter as well as sample group members would not pay tax but rather receive the transfer. Hence, an agenda–setter or sample group member of type z would receive a payoff of $zq_i + \tau$. Hence, τ is "sufficiently large" if it satisfies

$$\tau \geq z(\widehat{q}_i - q_i) - c(\widehat{q}_i).$$

[10] Voting proceeds in two stages. The first stage is used to select the alternative which is voted on in the second stage. To distinguish clearly between both stages, we use the term "selection" for the first stage, and the term "voting" for the second stage.

In a subclass of public good problems, transfers are not needed; that is, even $\tau = 0$ may satisfy the above inequality and be "sufficiently large." We discuss this issue in Subsection 5.3 and Corollary 1 below.

5.2 Discussion of the mechanism

We have shown that the PCSS mechanism implements the alternative that is the Condorcet winner. One may object, however, that the resulting *allocation* under our implementation result is not stable to majority voting: After all, the vast majority would prefer that the Condorcet–winning alternative be implemented without any tax–exemption of transfer to the agenda–setter and to sample group members. Since the sample group can be small, however, we can say that the PCSS mechanism implements the Condorcet winning alternative, while getting close to the Condorcet winning allocation. The small wedge between the implemented allocation and the Condorcet winning allocation is the price that one pays for accomplishing implementation of the Condorcet winning alternative. We emphasize that the mechanism only grants privileges conditional on the approval of the proposal by a super–majority of size $(1+\lambda)/2$. This is the same as saying that sample group members become privileged only if the majority of citizens outside the sample group give explicit approval to the proposal, and implicit approval to the privileges. This argument clarifies the motivation for the super–majority rule: Whenever a decision on a feasible alternative implies granting privileges to the sample group, this decision requires a majority of citizens outside the sample group. However, when citizens vote on a feasible alternative which is an amendment (rather than a proposal), no privileges are involved, and so a simple majority among the entire population suffices.

Compared to the baseline democratic mechanism, the mechanism discussed in the present section requires voting in two stages: First, there is a selection between the proposal, say q, and its "predecessor" $\psi_{-}(q)$. Second, the winner of that selection stage is pitted against the status quo. We stress that these two stages suffice for the implementation result, regardless of the number of states and alternatives. As an alternative decision–making procedure, one might have in mind to elicit preferences over all the pair of feasible alternatives. Clearly, this would require a number of voting stages increasing with $n^2 - n$. Thus, the democratic mechanism with sampling and two–stage voting is attractive from the point of view of procedural efficiency. In reality, it may often be costly to organize many voting stages, that is, many stages of referendum on the same issue. Therefore, procedural efficiency is an important concern, especially in large societies.

We note that the existence theorem does not depend on the prior belief of citizens about the probability distribution p on the states. All citizens are assumed to share a common prior belief about the probability of the different states. Our results do not depend on what exactly these prior probabilities are. This is an important and desirable robustness property of the mechanism.[11]

In all, we have seen that well–designed and conditional privileges for a small representative sample group may help societies overcome uncertainty about the underlying type distribution, share dispersed information, make more informed decisions, and choose policy options preferred by a majority.

We have argued that, in order to achieve implementation, it is vital to grant a privilege to some members of the society. This privilege consists of a tax–exemption and possibly a transfer. Taxes are typically not raised for one purpose at a time, nor to finance specifically one public good project. Therefore, when a citizen is granted a tax–exemption, this gives him a benefit which goes beyond not participating in the cost of the public good which we consider. In our model, such extra benefits would correspond to a cash transfer granted along with a tax–exemption. Therefore, a privilege which encompasses both a tax–exemption and a transfer seems well–motivated. Nevertheless, in the next subsection, we will show that implementation can be accomplished even if $\tau = 0$ under an extra restriction on the collective choice problem. Moreover, we will show that even without this restriction, implementation with $\tau = 0$ is possible if one is willing to add more selection stages. As a general remark on concerns about equal treatment and non–discrimination, we would like to stress that the privileges for sample group members could easily be thought of as a subsidization. In actual constitutions, equal treatment clauses do not apply to subsidies, and thus defining privileges as subsidies would avoid violating equal treatment principles (see e.g. Gersbach et al. (2015)).

Since sample group members are ordinary citizens, the practical implementation of necessary could take the following form: When filing their tax declarations, sample group members claim an exemption equal to the amount of tax that would be used for the specific decision that they have been part of. Moreover, they can demand that the overall tax bill is reduced by the amount of the transfer. Tax authorities can then verify whether the claims are justified. Of course, this is only one possible way to implement the scheme. Several alternatives are conceivable, too. For instance, sample group members could be subjected to regular taxation, but receive the amount of the tax exemption plus the transfer immediately after the decision is made.

5.3 Extensions

Theorem 6.1 states an implementation result provided that the transfer τ is sufficiently large. We have pointed out that this transfer can be suitably interpreted as the benefit of an exemption from all taxes other than the ones needed to finance the public good provision under consideration here. We are now going to derive a condition under which the implementation result holds even for $\tau = 0$.

[11] For the general theory of robust mechanisms in the standard framework, we refer to Bergemann and Morris (2005).

In the proof of Theorem 6.1, it turns out that a strictly positive transfer $\tau > 0$ is necessary for implementation only if there are a state $i \in N$, a type $z \in Z$, and a quantity $\widehat{q}_i \in Q$ which satisfy the inequalities

$$\widetilde{q}_i > \widehat{q}_i > q_i, \tag{6.1}$$
$$z\widehat{q}_i - c(\widehat{q}_i) > zq_i, \tag{6.2}$$

If such $i \in N$ and $\widehat{q}_i \in Q$ do not exist, then the implementation result of Theorem 6.1 holds even for $\tau = 0$. Thus, from the proof of Theorem 6.1, we obtain the following corollary:

Corollary 6.1. *Consider a subset of public good problems with the property that the type space Z, the cost function c, and the cumulative distribution functions $(F_i)_{i \in N}$ are such that for all $i \in N$ and for all $z \in Z$, it holds that*

$$z \leq \frac{c[\psi_-(\widetilde{q}_i)]}{\psi_-(\widetilde{q}_i) - q_i}.$$

On this subset of public good problems, the democratic mechanism with sampling and two–stage voting implements the Condorcet winner with $\tau = 0$.

In the description of the democratic mechanism with sampling and two–stage voting, we have imposed that any proposal $q \in Q$ be pitted against its "predecessor" $\psi_-(q)$ at the selection stage. This is equivalent to requiring the agenda–setter to propose a pair of two "successive" alternatives. Another way to accomplish implementation is to randomly appoint a second, "rival" agenda–setter who chooses the amendment and receives a reward (tax–exemption and transfer) if and only if the amendment is selected and becomes the final outcome of the mechanism.[12] In such a model, one can show that there is a Bayesian equilibrium in which the rival agenda–setter always believes with probability one that his proposal is the Condorcet winner and always suggests the predecessor of the proposal as an amendment. The existence of such an equilibrium follows from the same logic as the proof of Theorem 6.1.

One extension is to modify the democratic mechanism with sampling and selection by adding more selection stages. For this purpose, we define $\psi_-^t = \psi_-^{t-1}(\psi_-(q))$ for $t = 1, 2, \ldots$. Hence, in particular, $\psi_-^2(q) \equiv \max\{q \in Q | q < \psi_-(q)\}$. If the proposal is q, then the first selection stage would be between q and $\psi_-(q)$, the second stage between

[12] There is a strand of literature that deals with the case where two rival candidates for political office, rather than citizens, receive information about an underlying state of nature and the concomitant optimal policy, see for instance Heidhues and Lagerlöf (2003), Gratton (2014), Laslier and Van der Straeten (2014), and Kartik et al. (2015). In this branch of the literature, the crucial question is whether candidates for political office choose platforms in accordance with their information, or whether they "pander" to citizens' prior beliefs. This is in turn related to the issue of "pandering" in principal-agent relationships between a decision-maker and an expert, see for instance Che et al. (2013).

$\psi_-(q)$ and $\psi_-^2(q)$, and, in general, selection stage t between ψ_-^{t-1} and ψ_-^t, as long as an amendment defeats the current proposal. For such a more general mechanism with T selection stages, it can be shown that implementation of the Condorcet winner with $\tau = 0$ is possible if the type space Z, the cost function c, and the cumulative distribution functions $(F_i)_{i \in N}$ are such that for all $i \in N$ and for all $z \in Z$, it holds that

$$z \leq \frac{c[\psi_-^T(\widetilde{q}_i)]}{\psi_-^T(\widetilde{q}_i) - q_i}.$$

Thus, there is a trade–off between the mechanism's procedural efficiency and its ability to ensure equal treatment and, in particular, a limitation of the sample group's privileges.

One important extension of our model is to ask whether the implementation of the Condorcet winner can be accomplished if the status quo is not zero public good provision, but some other public good level $\bar{q} \in Q \setminus \{0\}$. If this is the case, then the implementation result of Theorem 6.1 satisfies a kind of "dynamic robustness" property: If the state of nature changes over time, then the democratic mechanism with sampling and two–stage voting can be applied again in order to move from the "old" Condorcet winner to the "new" Condorcet winner. We conjecture that, indeed, the democratic mechanism with sampling and two–stage voting can be suitably modified to implement the Condorcet winner for any given status quo. In particular, the democratic mechanism needs two features to ensure implementation:

- First, sample group members as well as the agenda–setter must be given a conditional privilege that is greater the more the agenda–setter's proposal differs from the status quo.

- Second, there must be a selection stage in which citizens choose between the proposal and an amendment which is slightly closer to the status quo than the proposal. As before, the sample group members' and agenda–setter's privilege must be conditional on the acceptance of the proposal.

To be more specific, one could define an (γ, θ)*–tax treatment* specifying that a citizen receives a lump sum transfer of θ and needs to pay a tax $\gamma c(q)$ when the public good level is $q \in Q$. If $\gamma > 1$ is chosen sufficiently high, then a citizen of any type $z \in Z$ who is subjected to an (γ, θ)–tax treatment strictly prefers zero public good provision over q_1, and strictly prefers q_i over q_{i+1} for any $i = 1, \ldots, n-1$. In addition, if $\theta > 0$ is chosen sufficiently high, then a citizen of any type $z \in Z$ prefers to be subjected to the (γ, θ)–tax treatment than to be a "regular" citizen who pays $c(q)$ and receives no transfer.

Consider the following modifications to the democratic mechanism with sampling and two–stage voting : At the selection stage, the proposal q is pitted against $\psi_-(q)$ (as before) if $q > \bar{q}$, but is pitted against $\psi_+(q)$ if $q < \bar{q}$. If $q = \bar{q}$, then $\bar{q}$ remains in effect. The agenda–setter and the sample group are granted a tax–exemption and receive the transfer of τ (as before) if the agenda–setter proposes some $q \in Q$ with $q > \bar{q}$, and subsequently q "wins" both at the selection and voting stages. The agenda–setter and the sample group are subjected to the (γ, θ)–tax treatment if the agenda–setter proposes some $q \in Q$ with $q < \bar{q}$, and subsequently q "wins" at both the selection stage and voting stage. As before, the special treatment of the sample group imposes a cost on each citizen outside the sample group but this cost vanishes in the limit when the mass of the sample group tends to zero.

6 Simplified implementation

So far in this paper, we have aimed at implementation for the entire set $\mathcal{P}$ of public good problems. In this section, we are going to restrict attention to the subset of public good problems which have the distance property. On that subset of public good problems, the agenda–setter's opportunities to exploit information sharing are restricted. This allows us to accomplish implementation of the Condorcet winner with a simpler mechanism.

We have previously defined the distance property. Intuitively, it means that feasible alternatives are drastically different from each other. When the distance property holds, any alternative higher than the Condorcet winner will lose against the status quo, and can therefore not be implemented. Thus, an agenda–setter who wants to exploit information sharing can only do so "in one direction," namely, by aiming at an alternative lower than the Condorcet winner. This allows for the Condorcet winner to be implemented using a simpler democratic mechanism. In particular, it requires only an unconditional privilege for a representative sample group. Moreover, a single voting stage suffices. More specifically, consider the following *democratic mechanism with sampling and one–stage voting.* A randomly drawn sample group of size $\lambda \in (0, 1)$ sends binary messages, and the information mapping φ as defined in the previous section is used to designate an alternative. Then, an agenda–setter chooses a proposal $q \in Q$, and citizens vote between q and the status quo. The agenda–setter and the sample group are tax–exempt.

Theorem 6.2. *For public good problems with the distance property, the Condorcet winner can be implemented by a democratic mechanism with sampling and one–stage voting, with $\lambda > 0$ sufficiently small.*

The proof can be found in Appendix D.

If the democratic mechanism with sampling and one–stage voting was applied to a public good problem without the distance property, it would lead to over–provision of public good. The reason is that the mechanism gives sample group members incentives to exaggerate the benefits from public good provision.

One might wonder which practical interpretation can be given to the distance property. Suppose that the issue at hand is to decide on a large infrastructure project, such as building a tunnel either with a single tube or with two tubes. This problem could be interpreted as having the distance property: There are few feasible alternatives, and their costs and benefits drastically differ from each other. As a counter–example, consider the problem of choosing the number of police officers in a city. This choice problem features many different alternatives among which one can choose "almost continuously." Such situations should be interpreted as not having the distance property.

7 Discussion, applications, and conclusion

The main insight of this paper is that democratic decision–making procedures can be used to identify and implement policies desired by the majority even in the presence of uncertainty about the type distribution. The resolution of such uncertainty and the associated implementation result hinge crucially on the use of a conditional privilege for a small sample group and the agenda–setter. The privilege itself and its conditional character motivate members of the sample not to *exploit* nor *manipulate* the resolution of uncertainty in any way.

Absent any such conditional privilege, however, democratic mechanisms based solely on communication prior to voting are prone to strategic behavior. In particular, there are two incentive problems: The agenda–setter's selfishness creates an incentive to *exploit* information sharing. Even in the absence of exploitation, information sharing is prone to *manipulation.*

We stress that the introduced democratic mechanisms do not depend on the citizens' ex ante beliefs about the states of nature. This is a particularly desirable robustness property of democratic mechanisms, since these mechanisms should be applicable to a variety of situations and since their rules should not depend on citizens' current beliefs.

Our paper suggests a three–stage procedure that reveals and implement the Condorcet winner when one player acts as agenda–setter. We ask under what conditions citizens have incentives to share their private information and thereby resolve uncertainty about the underlying type distribution. The paper does, however, also allow for some conclusions about settings without such uncertainty about the underlying type distribution. To see this, suppose that citizens were informed about the underlying distribution of preferences in the population. In that case, the communication stage of our democratic mechanism would be redundant. However, the selection stage of the democratic mechanism would still be useful as a safeguard against the selfishness of an agenda–setter. For instance, think of the agenda–setter as a bureaucrat who wishes to maximize the size of public projects. It is a well–known theme in the political economy literature that sometimes privileged groups, such as members of the government bureaucracy, would like to maximize public good provision in order to secure their own perks or ego rents. One could explore to what extent the corrective present in the selection stage of our democratic mechanism with sampling and two–stage voting and the related conditional character of the privileges could be applied to that problem. Giving the citizens a vote between the proposed alternative and a "smaller" alternative could protect the public from this kind of behavior in some circumstances (depending how different the alternatives are from each other).

In the present paper, we assume that the message space of all citizens accept the agenda–setter is binary. This restriction applies both at the communication and voting stages. In the latter case, the motivation is that a referendum typically reduces complex decisions to a matter of simple Yes–or–No approval.

At the communication stage, the restriction to binary messages is a mere simplification. Our results would persist if we allowed for a richer message space at the communication stage. This is straightforward with regard to our existence result: Implementation is possible with binary messages, and citizens with richer message spaces can replicate binary messages, so the existence result holds for richer message spaces as well.

For the impossibility result, what we need is that a particular group of citizens can emulate in one state the messages that would be sent in another state. As long as all citizens have the same finite message space, this remains true, and so the impossibility result holds, even if the message space is not binary.

There is a variety of extensions and further applications which can be considered in future research. For instance, one could examine to what extent our results carry over to choices from different sets of possible policies, such as continuous policy spaces, or multi–dimensional collective choice problems in which several public goods can be combined in a bundle of public goods. Moreover, one might consider an electorate with different income levels and the possibility to differentiate the tax burden as a function of income. In such a model, one could investigate the effect of a policy chosen by a democratic mechanism on the degree of inequality among citizens.

Of course, an alternative way to implement the Condorcet winner is to have a social planner ask all citizens at once for their entire preference ordering. Such a process, however, is prone to manipulation by citizens. There is a large literature on this topic starting with the famous Gibbard–Satterthwaite theorem. One may wonder if citizens could not be asked to submit their entire preference profile on a single ballot paper. This method may not be feasible in reality if the set of alternatives is too large. There are, however, real world examples of ballots in which citizens express preferences over three alternatives. Such three–way ballots have been applied in the Swiss system of direct democracy. In our model, three–way ballots could be used to merge the selection stage and the voting stage of the mechanism, which would further enhance its procedural efficiency.

Another relevant question is how the mechanisms discussed in the present paper could be applied in parliamentary settings. Procedural efficiency is an issue in parliamentary democracy as well, since excessively long deliberations on a single issue have opportunity costs, and distract attention from other topics. Procedurally efficient parliamentary decisions might be achieved by a democratic mechanism with sampling and two–stage voting with a randomly chosen sample group from the population and a subsequent decision by a majority in parliament.

While such extensions will considerably expand the scope of democratic mechanisms in a polity, it is likely that optimal democratic mechanisms with procedural efficiency will involve conditional tax privileges for small groups. In the presence of uncertainty about the type distribution, we expect such conditional privileges to be an essential ingredient of democratic mechanisms.

Appendix A

In Appendix A, we provide a detailed account of assumptions on the cost function, the probability distributions, the type space, and the Condorcet winners which generate the properties imposed in Section 2 and, in particular, in Assumption 1.

Cost function

Although we will eventually consider a discrete set of possible quantities, it is convenient to start by defining continuous cost and utility functions on $\mathbb{R}_+$. In particular, we assume that the per capita cost of providing a quantity $q \in \mathbb{R}_+$ of the public good is given by a twice continuously differentiable, strictly increasing, and strictly convex function $c : \mathbb{R}_+ \to \mathbb{R}_+$ with $c(0) = 0$. This implies in particular that average cost $c(q)/q$ is strictly increasing. One typical interpretation of such a cost function is that each citizen is initially endowed with w ($w > 0$) units of a private consumption good which can either be consumed or transformed into the public good. The per capita costs $c(q)$ then represent the utility losses due to foregone private consumption.

Probability distributions

We assume that the family of probability distributions from which the types are drawn in each state can be represented by continuously differentiable probability density functions which satisfy a property known as the *monotonicity of likelihood ratios.* More formally, let the first derivative of f_k be $(f'_k)_{k=1,\dots,n}$. Then, we assume that $f_k(z) > 0$ for all $k \in N$ and all $z \in Z$ and, moreover, that

$$\frac{f'_{k+1}(z)}{f_{k+1}(z)} > \frac{f'_k(z)}{f_k(z)}, \ \forall z \in Z, \ \forall k \in N \setminus \{n\}. \tag{6.3}$$

The monotonicity of likelihood ratios property has three key implications. First, the probability distribution associated with F_{k+1} (strictly) first–order stochastically dominates the one associated with F_k for every $k \in N \setminus \{n\}$, that is, $F_{k+1}(z) < F_k(z)$ for every $z \in int(Z)$. In that sense, the benefits from the public good are higher in state $k+1$ than in state k. Second, the monotonicity of likelihood ratios implies a single–crossing property of the probability density functions, which will be crucial for our analysis. Finally, the monotonicity of likelihood ratios imposes monotone Bayesian updating (Milgrom 1981). More specifically, citizen z believes in state $k = 1, \dots, n$ with probability

$$\beta_k(z) = \frac{f_k(z)p_k}{\sum_{j=1}^n f_j(z)p_j}, \tag{6.4}$$

and the associated probability distributions for citizen z_2 stochastically dominates the one for citizen z_1 if $z_1 < z_2$.

We note that the monotonicity of likelihood ratios implies, loosely speaking, that a higher type tends to believe in higher states of nature with higher probability.

Type space

We assume that, for every $z \in Z$, the most desired public good level of citizen z belongs to Q. Formally, since the most preferred public good level of citizen z is given by $z = c'(q)$, we assume $c'(0) \leq inf(Z)$ and $c'(q^{max}) \geq sup(Z)$ where $inf(Z)$ and $sup(Z)$ are the infimum and supremum of the type space Z and q^{max} the highest possible public good level in Q.[13]

Together with the strict convexity of the cost function, this property implies that the preferences of each type $z \in Z$ are single–peaked, and the most preferred quantity of citizen z is that q which solves $c'(q) = z$.

Moreover, we assume $c(q_1)/q_1 \in int(Z)$.

Condorcet winner

Citizen z is the *median voter* in state k if $F_k(z) = 1/2$. We define for each state $k \in N$ a quantity $q_k \in \mathbb{R}_+$ as the unique solution to

$$F_k\left(c'(q_k)\right) = 1/2,$$

so that q_k is the most preferred quantity of the median voter in state k. The definition and the assumptions regarding the cost function and the monotonicity of likelihood ratios imply that $0 < q_1 < \ldots < q_n$. Due to the strict convexity of the cost function, we have

$$\begin{aligned} \frac{c(q_k) - c(q)}{q_k - q} &< c'(q_k),\ \forall q < q_k,\ \forall k \in N, \\ \frac{c(q_k) - c(q)}{q_k - q} &> c'(q_k),\ \forall q > q_k,\ \forall k \in N. \end{aligned}$$

Due to the single–peaked preferences, these statements can be expressed in terms of the distribution functions as Inequalities (6.5) and (6.6) below. These expressions are well–defined since $c(q_1)/q_1 \in int(Z)$.

$$F_k\left(\frac{c(q_k) - c(q)}{q_k - q}\right) < 1/2,\ \forall q < q_k,\ \forall k \in N, \tag{6.5}$$

$$F_k\left(\frac{c(q_k) - c(q)}{q_k - q}\right) > 1/2,\ \forall q > q_k,\ \forall k \in N. \tag{6.6}$$

Verbally, in state k, a simple majority of citizens prefers q_k over any other quantity $q \in \mathbb{R}_+ \setminus \{q_k\}$. Thus, q_k is the *Condorcet winner* in state k.

Point of indifference

[13] We assume here that there is a maximal element of Q. This is not essential, however. If one wants to allow for an unbounded set Q, one would have to assume instead that the second derivative $c''(q)$ is strictly positive and bounded away from zero.

Moreover, for every $k \in N$, we define the quantity $\widetilde{q}_k \in \mathbb{R}_+$ as the unique[14] solution to

$$F_k\left(c(\widetilde{q}_k)/\widetilde{q}_k\right) = 1/2. \tag{6.7}$$

In state k, a majority of citizens prefers any quantity $q < \widetilde{q}_k$ to zero public good, but prefers zero public good to any quantity $q > \widetilde{q}_k$. Given that average cost $c(q)/q$ is strictly increasing and that $c(q)/q < c'(q)$, we have $\widetilde{q}_k > q_k$ for every $k \in N$.

Appendix B

Proof of Proposition 6.1

Take any $P \in \mathcal{P}$, and suppose that citizens do not manipulate information sharing. That is, they use a communication strategy σ^P such that $\varphi(\delta_i(\sigma^P)) = q_i$ for every $i \in N$. Due to sincere voting, if the true state is i, the agenda–setter anticipates that any proposal $q \in Q$ such that $q < \widetilde{q}_i$ is going to win against the status quo. Suppose that the agenda–setter's type z is such that $z = c'(q)$ for some $q \in Q$ such that $q < \widetilde{q}_i$ and $q \neq q_i$. In this case, it is in the agenda–setter's interest to choose $\rho^P(z, \delta_i(\sigma^P)) = q$ rather than $\rho^P(z, \delta_i(\sigma^P)) = q_i$. Indeed, he can exploit information sharing. □

Proof of Proposition 6.2

Suppose by way of contradiction that a baseline democratic mechanism implements the Condorcet winner. Take some public good problem $P \in \mathcal{P}$ which has the distance property, and in which a state $i \in N \setminus \{n\}$ is concealable. Let σ^P be a communication strategy and ρ^P a proposal strategy such that

$$\varphi^P(\delta_k(\sigma^P)) = \rho^P(z, \delta_k(\sigma^P)) = q_k$$

for every $k \in N$ and every $z \in Z$. By construction of φ^P, the communication strategy σ^P is such that citizen z votes Yes if and only if $z \geq z^P$. We show that σ^P cannot be optimal.

Due to the premise that state $i \in N \setminus \{n\}$ is concealable, the function $F_i(z) - F_{i+1}(z)$ has a unique maximizer z_i^* which belongs to Z_-.

Let $\widetilde{z} = \min\{z^P, z_i^*\}$. We have

$$\delta_{i+1}(\sigma^P) - \delta_i(\sigma^P) \leq F_i(\widetilde{z}) - F_{i+1}(\widetilde{z}).$$

In particular, this inequality implies

$$\delta_i(\sigma^P) + F_i(\widetilde{z}) \geq \delta_{i+1}(\sigma^P).$$

[14]The uniqueness of $\widetilde{q}_k$ follows again from the strict convexity of the cost function.

Due to continuity of the cumulative distribution function, there is a ζ such that $\zeta \leq \widetilde{z} \leq z_i^* \leq \widehat{z}$ and

$$\delta_i(\sigma^P) + F_i(\zeta) = \delta_{i+1}(\sigma^P).$$

Consider a joint deviation from σ^P, under which citizens $z \leq \zeta$ vote Yes, and citizens $z > \zeta$ vote according to σ^P. Since $\zeta \leq \widehat{z}$, this is a deviation by members of Z_- only. If the true state is i, then the resulting share of Yes–votes is $\delta_i(\sigma^P) + F_i(\zeta) = \delta_{i+1}(\sigma^P)$. Thus, under this deviation, the proposal q_{i+1} will be made at the voting stage with probability one if the true state is i. Since citizens vote sincerely, this proposal will be rejected since $q_{i+1} > \widetilde{q}_i$, and the resulting public good level will be zero. By construction, members of Z_- prefer zero over q_i. Now we have shown that the deviation by Z_- is (strictly) profitable if the true state is i. Suppose next that the true state is some $j \in N \setminus \{i\}$. Due to the distance property, no quantity $q \in Q$ with $q > q_j$ would be accepted at the voting stage in state j. Thus, under the deviation, the outcome must be some $q'_j \leq q_j$ if the state is j. All members of Z_- weakly prefer q'_j to q_j. Since $p_i > 0$ and thus $\beta_i(z) > 0$ for all $z \in Z$, the deviation is (strictly) profitable in expectation. □

Appendix C

Proof of Theorem 6.1

Now we turn to the proof of the implementation result. First, we are going to consider the final voting stage. The criterion for sincere voting differs depending on whether the final vote is between the proposal and the status quo or between the amendment and the status quo. We state the appropriate definition of sincere voting. Second, we are going to establish conditions under which decisions at the selection stage are also sincere, in a sense to be made precise. Third, we state a profile of strategies and beliefs which is a Bayesian equilibrium, involves sincere selection, and leads to the implementation of the Condorcet winner.

Sincere voting

The voting stage of the democratic mechanism with sampling and two–stage voting is a final and binary choice which leaves no room for profitable strategic behavior. As a result, citizens vote sincerely at the voting stage. However, the sample group receives a privilege conditional on final acceptance of the proposal. This conditional privilege must be taken into account in order to determine the appropriate meaning of "sincere voting" in the present context. Suppose that at the voting stage, the alternative $q \in Q$ is pitted against the status quo, which is zero public good provision. Consider first the case where q is not the original proposal, but the amendment. In that case, neither tax–exemptions nor transfers will be granted, so that sincere voting simply means that citizen z votes for q if and only if $z \geq c(q)/q$.

Now suppose instead that q is the original proposal made by the agenda–setter. Since $zq > 0$ for all $q \in Q \setminus \{0\}$ and all $z \in Z$, it is sincere for every sample group member to vote in favor of q. Citizen z, who is not a sample group member, votes sincerely in favor of q if

$$z > \frac{c(q) + \lambda\tau}{(1-\lambda)q}.$$

The conditional privilege distorts sincere voting behavior for citizens outside the sample group. We argue that this distortion is negligible, however, when $\lambda > 0$ is small enough.

For any given values of λ and τ, and for every state $i \in N$, define $\widetilde{q}_i^{(\lambda,\tau)}$ as the solution to the equality

$$F_i\left(\frac{c(\widetilde{q}_i^{(\lambda,\tau)}) + \lambda\tau}{(1-\lambda)\widetilde{q}_i^{(\lambda,\tau)}}\right) = 1/2.$$

We note that $\widetilde{q}_i^{(\lambda,\tau)} < \widetilde{q}_i$ for $\lambda > 0$ and $\tau \geq 0$. Moreover, for any value of $\tau \geq 0$, we find that $\widetilde{q}_i^{(\lambda,\tau)}$ converges to $\widetilde{q}_i$ as $\lambda \downarrow 0$. Verbally, considering the limit as the sample group becomes arbitrarily small, the distortion of sincere voting behavior due to the conditional privilege vanishes.

Sincere selection

We turn to the selection stage which precedes the voting stage. In this selection stage, voters select either a proposal q or an amendment $\psi_-(q)$. In principle, a "selection strategy" should specify whether a citizen of a given type with a given belief about the state of nature selects any given proposal or the associated amendment. For the purpose of our argument, however, it will be enough to specify the optimal selection by citizens who believe with probability one that any alternative selected at the selection stage will win against the status quo at the voting stage. In that case, the decision to select the proposal or the amendment is a final and binary choice of the alternative to be implemented. Consequently, it is optimal to select sincerely, and we only have to spell out what sincere selection means: A citizen z outside the sample group prefers proposal q over amendment $\psi_-(q)$ if $zq - \frac{c(q)+\lambda\tau}{(1-\lambda)} > z\psi_-(q) - c(\psi_-(q))$. A sample group member always prefers q over $\psi_-(q)$ because $\tau + zq > z\psi_-(q) - c(\psi_-(q))$, that is, a sample group member sincerely selects the proposal.

Strategies and beliefs

We now proceed with the construction of a communication strategy, a proposal strategy, and some consistent beliefs. The communication strategy is as follows:

$$\sigma^P(z) = \begin{cases} Yes & \text{if } \ z \geq z^P, \\ No & \text{otherwise.} \end{cases}$$

The proposal strategy is such that the agenda–setter does not exploit information sharing, that is, for every $z \in Z$, we have $\rho^P(z, \delta) = \varphi^P(\delta)$, thus:

$$\rho^P(z, \delta) = \begin{cases} q_1 & \text{if } 0 \leq \delta < 1 - F_2(z^P), \\ q_k & \text{if } 1 - F_k(z^P) \leq \delta < 1 - F_{k+1}(z^P), \ k = 2, \ldots, n-1, \\ q_n & \text{if } 1 - F_n(z^P) \leq \delta \leq 1. \end{cases}$$

The agenda–setter (regardless of his type) believes with probability one that information sharing has been "successful" so that the alternative identified by the information mapping is the Condorcet winner. Thus, using e_i to denote the n–vector with component i equal to one and all other components equal to zero, we can write his beliefs as follows:

$$\pi^P_{AS}(\delta) = \begin{cases} e_1 & \text{if } 0 \leq \delta < 1 - F_2(z^P), \\ e_k & \text{if } 1 - F_k(z^P) \leq \delta < 1 - F_{k+1}(z^P), \ k = 2, \ldots, n-1, \\ e_n & \text{if } 1 - F_n(z^P) \leq \delta \leq 1. \end{cases}$$

Citizens other than the agenda–setter, regardless of their type, believe with probability one that the proposal $q \in Q$ made by the agenda–setter is the Condorcet winner, thus:

$$\pi^P(q) = \begin{cases} e_1 & \text{if } q < q_2, \\ e_k & \text{if } q_k \leq q < q_{k+1}, \ k = 2, \ldots, n-1, \\ e_n & \text{if } q \geq q_n. \end{cases}$$

We note that citizens other than the agenda–setter base their beliefs only on the proposal q made by the agenda–setter, while the observed share δ of positive signals determines the agenda–setter's beliefs. This construction of beliefs has allowed us to specify optimal selection only for the case where citizens believe that whichever alternative they select at the selection stage is going to become the final outcome of the mechanism.

Implementation

In order to establish the theorem, we need to show two things: First, the communication strategy σ^P and the proposal strategy ρ^P as defined above actually lead to the implementation of the Condorcet winner in every state. Second, these strategies are optimal given the beliefs $\pi^P_{AS}(\delta)$ and $\pi^P(q)$.

Implementation.
We see that $\rho^P(z, \delta_k(\sigma^P)) = \varphi^P(\delta_k(\sigma^P)) = q_k$ for every $k \in N$. Thus, if the communication strategy σ^P and the proposal strategy ρ^P are played, then the agenda–setter believes with probability one in the true state, and he proposes the Condorcet winner. The beliefs π^P_{AS} and π^P are such that citizens believe with probability one that the quantity proposed by the agenda–setter is the Condorcet winner. Hence, all citizens believe that both the proposal made by the agenda–setter and the amendment would prevail against the status quo at the voting stage. Citizens select sincerely at the selection stage. In particular, all sample group members select the proposal. Since the proposal is the Condorcet winner, also at least half of the citizens outside the sample group select it. Hence, in the entire population, a share of at least $(1 + \lambda)/2$ select the proposal. Because of sincere voting at the voting stage and the fact that the proposal is the Condorcet winner on the path of play, the Condorcet winner does become the final outcome of the mechanism, as desired.

Optimal signaling and agenda–setting.
Now we show that no deviation by the agenda–setter or by sample group members can be profitable. Note that sample group members and agenda–setter receive the payoff $zq_k + \tau \geq zq_k$ on the path induced by the profile under consideration when the state is k. Suppose by way of contradiction that there is a profitable deviation for the agenda–setter or (some coalition within) the sample group. For every $k \in N$, let $\widehat{q}_k$ be the outcome which follows after a profitable deviation. Due to the premise that the deviation is profitable, there must be a state $i \in N$ such that $\widehat{q}_i > q_i$. This is only possible if the agenda–setter proposed either $\widehat{q}_i$ or $\psi_+(\widehat{q}_i)$. Either way, by construction of π^P, citizens believe with probability one that both the proposal and the amendment would win against the status quo at the voting stage. Hence, citizens outside the sample group select sincerely, while sample group members select the proposal. However, the majority of citizens outside the sample group prefer q_i over $\widehat{q}_i$ simply because q_i is the Condorcet winner in state i. Hence, $\widehat{q}_i > q_i$ can only be the outcome of the mechanism if it has been an amendment at the selection stage, and has then prevailed against the status quo $q = 0$ at the voting stage. Hence, under the deviation, there is no tax exemption and no transfer. Since the deviation is profitable for an agenda–setter or sample group member of type z, it must hold that $z\widehat{q}_i - c(\widehat{q}_i) > zq_i + \tau$. However, for any $z \in Z$, this inequality is violated for sufficiently large τ. Indeed, we obtain a contradiction for large enough τ. $\square$

Appendix D

In Appendix D, we show that the democratic mechanism with sampling and one–stage voting implements the Condorcet winner in those public good problems which satisfy the distance property.

Sincere voting

From the point of view of citizens outside the sample group, the provision of a public good quantity $q \in Q$ is no longer associated with a tax burden of $c(q)$, but of $\left(\frac{1}{1-\lambda}\right) c(q)$. Consequently, the utility of citizen z outside the sample group from the public good level $q \in Q$ is given by

$$\widehat{u}(z, q) = zq - \left(\frac{1}{1-\lambda}\right) c(q).$$

The popular vote in the democratic mechanism with sampling is a binary and final choice and thus strategic voting can never be beneficial. The citizens vote sincerely in the democratic mechanism with sampling and one–stage voting. We now formally spell out the sincere voting rule in the democratic mechanism with sampling and one–stage voting. Of course, this rule is different for sample group members and for regular citizens. Indeed, suppose that at the voting stage of the democratic mechanism with sampling and one–stage voting the alternative $q \in Q$ is pitted against the status quo, which is zero public good provision. Since $Z \subset \mathbb{R}_{++}$, we have that $zq > 0$ for every $q \in Q \setminus \{0\}$, and thus all sample group members vote sincerely in favor of the proposal. Citizen z, who is not a sample group member, prefers q to zero public good provision if

$$z > \frac{c(q)}{(1-\lambda)q}.$$

Analogously to the argument in Appendix C, the privileges for the sample group distort sincere voting behavior of citizens outside the sample group. Again, we argue that this distortion is negligible when λ is small.

For every state $i \in N$, define $\widetilde{q}_i^\lambda$ as the solution to the equality

$$F_i\left(\frac{c(\widetilde{q}_i^\lambda)}{(1-\lambda)\widetilde{q}_i^\lambda}\right) = 1/2.$$

In words, if the true state is i, then a majority of citizens outside the sample group prefers any quantity $q \in Q$ with $q < \widetilde{q}_i^\lambda$ over zero public good provision, and prefers zero public good provision to any quantity $q \in Q$ such that $q > \widetilde{q}_i^\lambda$. The assumptions introduced on the cumulative distribution functions $(F_i)_{i \in N}$ guarantee that $\widetilde{q}_i^\lambda$ exists.

For every $\lambda > 0$, we have $\widetilde{q}_i^\lambda < \widetilde{q}_i$, and we find that $\widetilde{q}_i^\lambda$ converges to $\widetilde{q}_i$ as $\lambda \downarrow 0$. Suppose that $q \in Q$ is a proposal which is preferred by the majority to zero public good provision under uniform taxation. Then, q is also preferred to zero public good provision by the majority under the tax–exemption, provided that $\lambda > 0$ is small enough.

Strategies and beliefs

Define the following communication strategy σ^P:

$$\sigma^P(z) = \begin{cases} Yes & \text{if } \ z \geq z^P, \\ No & \text{otherwise.} \end{cases}$$

Moreover, consider the proposal strategy:

$$\rho^P(z,\delta) = \begin{cases} q_1 & \text{if } 0 \leq \delta < 1 - F_2(z^P), \\ q_k & \text{if } 1 - F_k(z^P) \leq \delta < 1 - F_{k+1}(z^P), \ k = 2, \ldots, n-1, \\ q_n & \text{if } 1 - F_n(z^P) \leq \delta \leq 1. \end{cases}$$

Define a belief for the agenda–setter as:

$$\pi^P_{AS}(\delta) = \begin{cases} e_1 & \text{if } 0 \leq \delta < 1 - F_2(z^P), \\ e_k & \text{if } 1 - F_k(z^P) \leq \delta < 1 - F_{k+1}(z^P), \ k = 2, \ldots, n-1, \\ e_n & \text{if } 1 - F_n(z^P) \leq \delta \leq 1. \end{cases}$$

It is straightforward that the belief π^P_{AS} is consistent with the strategies σ^P and ρ^P. Moreover, if these strategies are played, then the outcome is the Condorcet winner. In order to establish the implementation result, we need to show that σ^P and ρ^P are optimal given the beliefs π^P_{AS}. We show first that the proposal strategy ρ^P is optimal. Indeed, let i be the true state, so that q_i is the Condorcet winner. Due to the distance property and sincere voting, any proposal $q \in Q$ such that $q > q_i$ will be rejected, and any proposal $q \in Q$ such that $q \leq q_i$ will be accepted in the popular vote. Since the agenda–setter is tax–exempt, it follows immediately that it is optimal for him to propose the quantity q_i whenever his belief is e_i. Otherwise, if the agenda–setter makes any proposal $q \in Q \setminus \{q_i\}$, the outcome of the mechanism is some $q' \in \{q, 0\}$, and $q' < q_i$. Next we show that the communication strategy is optimal. All members of the sample group as well as the agenda–setter are tax–exempt. A joint deviation by a subset of the sample group (or a deviation by the agenda–setter) could only be profitable if it led to some public good level $q > q_i$. Due to sincere voting, however, such a public good level will not be implemented.

□

References

Aghion, P. and P. Bolton (2003) "Incomplete Social Contracts," *Journal of the European Economic Association, 1,* 38-67.

Aidt, T. and F. Giovannoni (2011) "Critical Decisions and Constitutional Rules," *Social Choice and Welfare, 37(2),* 219-268.

Bednar, J. (2011) "Nudging Federalism Towards Productive Experimentation," *Regional and Federal Studies, 4-5,* 503-521.

Bowen, H. (1943) "The Interpretation of Voting in the Allocation of Economic Resources," *Quarterly Journal of Economics, 58,* 27–48.

Bendor, J. and D.Mookherjee (1987) "Institutional Structure and the Logic of Ongoing Collective Action," *American Political Science Review, 81,* 129-154.

Bergemann, D. and S.Morris (2005) "Robust Mechanism Design," *Econometrica, 73,* 1771-1813.

Bernhardt, D., J. Duggan, and F.Squintani (2008) "A Survey on Polling in Elections," in Aragones, E., Bevia, C., Llavador, H., Schofield, N., Eds., *The Political Economy of Democracy*, Fundacion BBVA.

Bierbrauer, F.J. and M.F. Hellwig (2016) "Robustly Coalition–proof Incentive Mechanisms for Public Good Provision are Voting Mechanisms and vice versa," *Review of Economic Studies, 83,* 1440–1464.

Bouton, L. (2013) "A Theory of Strategic Voting in Runoff Elections," *American Economic Review, 103,* 1248–88.

Britz, V. and H.Gersbach (2014) "Experimentation in Democratic mechanisms," Center of Economic Research at ETH Zurich Working Paper No. 14/199.

Buchanan, J.M. and G.Tullock (1962) *The Calculus of Consent: Logical Foundations of Constitutional Democracy,* Ann Arbor; MI: University of Michigan Press.

Buera, F.J., A.Monge-Naranjo, and G.E.Primiceri (2011) "Learning the Wealth of Nations," *Econometrica, 79,* 1-45.

Cai, H. and D.Treisman (2009) "Political Decentralization and Policy Experimentation," *Quarterly Journal of Political Science, 4(1),* 35-58.

Callander, S. and B.Harstad (2013) "Experimentation in Federal Systems," NBER Working Paper 19601.

Che, Y.K., W. Dessein, and N. Kartik (2013) "Pandering to Persuade," *American Economic Review, 103,* 47–79.

Cho, S.J. (2014) Voting Equilibria under Proportional Representation, *American Political Science Review, 108,* pp

Crawford, V.P. and J.Sobel (1982) "Strategic Information Transmission," *Econometrica, 50,* 1431-1451.

Dellis, A. (2013) "The two Party System under Alternative Voting Procedure," *Social Choice and Welfare, 40,* 263–284.

Gersbach, H. (2009) "Democratic Mechanisms," *Journal of the European Economic Association, 7,* 1436-1469.

Gersbach, H. (2017) "Flexible Majority Rules in Democracyville: A Guided Tour," *Mathematical Social Sciences, 85,* 37–43.

Gersbach, H., V. Hahn, and S.Imhof (2013) "Tax Rules," *Social Choice and Welfare, 41,* 19-42.

Gershkov, A., B. Moldovanu, and Xianwen Shi (2017) "Optimal Voting Rules," *Review of Economic Studies, 84,* 688–717.

Gratton, G. (2014) "Pandering and Electoral Competition," *Games and Economic Behavior, 84,* 163–179.

Heidhues, P. and J. Lagerlöf (2003) "Hiding Information in Electoral Competition," *Games and Economic Behavior, 42,* 48–74.

Kartik, N., F. Squintani, and K. Tinn (2015) "Information Revelation and Pandering in Elections." Mimeo.

Kollman, K. (2003) "The Rotating Presidency of the European Council as a Search for Good Policies," *European Union Politics, 4,* 51-74.

Kollman, K., J.H.Miller, and S.E.Page (2000) "Decentralization and the Search for Policy Solutions," *Journal of Law, Economics, and Organization, 16,* 102-128.

Laslier, J.F. and K. Van der Straeten (2004) "Electoral Competition under Imperfect Information," *Economic Theory, 24,* 419–446.

Meirowitz, A. (2003) "Polling Games and Information Revelation in the Downsian Framework," *Games and Economic Behavior, 51,* 464-489.

Milgrom, P.R. (1981) "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics, 12,* 380–391.

Morgan, J. and P.C. Stocken (2008) "Information Aggregation in Polls," *American Economic Review, 98,* 864-896.

Rasch, B.E. (2014) "Insincere Voting under the Successive Procedure," *Public Choice, 158,* 499–511.

Rose-Ackerman, S. (1980) "Risk-taking and Reflection: Does Federalism Promote Innovations?" *Journal of Legal Studies, 9(3),* 593–616.

Shipan, C.R. and C.Volden (2006) "Bottom–up Federalism: The Diffusion of Anti–smoking Policies from U.S. Cities to States," *American Journal of Political Science, 50,* 825-843.

Strumpf, K.S. (2002) "Does Government Decentralization Increase Policy Innovation?" *Journal of Public Economic Theory, 4,* 207-241.

Van der Straeten, K., J.F. Laslier, N., Sauger, A. Blais (2010) " Strategic, Sincere, and Heuristic Voting under Four Election Rules: An Experimental Study, *Social Choice and Welfare, 35,* 435–472.

Volden, C. (2006) "States as Policy Laboratories: Emulating Success in the Children's Health Insurance Program," *American Journal of Political Science, 50,* 294-312.

Volden, C., M.M.Ting, and D.P.Carpenter (2008) "A Formal Model of Learning and Policy Diffusion," *American Political Science Review, 102,* 319-332.