HEIDELBERG UNIVERSITY
DEPARTMENT OF ECONOMICS

UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

The “German Vote” and its consequences: (Un)reliable parties in multilateral bargaining under private information

David Piazolo

AWI DISCUSSION PAPER SERIES NO. 756
October 2024

The "German Vote" and its consequences: (Un)reliable parties in multilateral bargaining under private information

David Piazolo[a]

[a]*Heidelberg University, Alfred Weber Institute for Economics, Bergheimer Straße 58, Heidelberg, 69115, Germany*

Abstract

This paper theoretically investigates the strategic implications of varying *reliability* of bargaining partners under unanimous and non-unanimous voting. Three players (one proposer, two responders) bargain over the distribution of a pie. One responder has private information about his valuation of finding an agreement, implying signaling values that differ substantially between voting rules and are affected by the other responder's reliability. Under unanimity rule, the responder with private information benefits from voting "no" because this signals that he requires a larger compensation in a future period. In contrast, under majority rule, voting "no" is unattractive due to the *fear of being excluded* from a future coalition. Under both voting rules, one responder becoming less reliable negatively affects the other responder's willingness to vote "yes", making efficient agreements increasingly difficult to achieve. The presence of unreliable parties can under majority rule lead to more parties being included in the winning coalition, as demonstrated by an extension of the model.

Keywords: bargaining, majority, unanimity, unreliability, private information
JEL: C78, D72, P48

Email address: `david.piazolo@awi.uni-heidelberg.de` (David Piazolo)

Preprint submitted to Elsevier *October 22, 2024*

1. Introduction

> "A German abstention during the vote on the EU supply chains law would send a terrible political message [...] and it would jeopardise Germany's credibility — after all, German political leaders [...] have previously agreed that the EU supply chain law should be adopted, have actively engaged in negotiations, and shaped the most recent draft. There is also a risk that if Germany abstains, other countries will reconsider their position and withdraw their support, creating a domino effect." (Human Rights Watch (2024))

Lately, the "German Vote" has become a term being discussed extensively at the EU Council in Brussels. Germany twice halted the introduction of a new law in a last minute blockade after having signalled being in favor of the new law in the months before. Such line of action is highly unusual and not well-respected in Brussels. Still, bargaining partners becoming increasingly unreliable appears to be a phenomenon prevalent around the world, be it the US with the possible return of Donald Trump into the White House, authoritarian states with unpredictable leaders, or, as mentioned above, Germany in the EU Council due to dissent between domestic ruling parties. More generally, any bargaining partner facing domestic elections tends to be somewhat unreliable, as a new government might have completely different views on the discussed issues.

This paper aims at modeling the respective consequences of having unreliable bargaining partners under the application of both unanimity and simple majority voting rules. Suppose France (the Proposer) plans the introduction of a new law at the EU level and therefore needs to convince the other states, say Germany and Italy (the responders), to vote yes. While France benefits from the introduction of the new law by one unit, the other states have private information about their valuation of the law. As I will demonstrate in a first step, a responder under unanimity rule can signal their dislike for the agreement by voting No, thereby indicating that they require substantial compensation to be convinced to vote yes. In contrast, under majority rule, signaling to dislike agreement by a No-vote leads to an increased risk of being excluded from a future coalition. Hence responders become expensive under unanimity rule whereas an agreement is much easier to be reached under majority voting. Comparing the effects under various voting rules is crucial because the voting rules applied in the European Council vary across different policy areas.[1]

[1] At the EU Council, there is an ongoing debate whether to apply the "qualified majority rule",

I focus on the interaction between the Proposer and one of the responders (R1, say Italy), while assuming that the second responder (R2, say Germany) follows an exogenously fixed voting rule which varies in its "reliability". Specifically, I assume that a fully "reliable" R2 votes yes with certainty whenever he is offered more than a certain exogenously fixed share c. In contrast, an "unreliable" R2 votes no with some probability p even when offered more than c. Under unanimity rule, this resembles the bilateral bargaining situation as in Schnedler and Vanberg (2014). Private information about one's valuation of agreement leads to their idea of "playing hard to get": Sophisticated players act as if they dislike a task in order to be compensated for it in the future.

This perfectly corresponds to the so-called *positive signaling value from voting No* in my model. By voting yes, responders, similarly, signal that they are "easy to get" (i.e., "cheap"). Italy must fear such *negative signaling value from voting yes* only in case a yes-vote does not secure immediate agreement with certainty. This is indeed the case when Germany is an unreliable bargaining partner who might vote no even though being offered the claimed share of c. As a result, a yes-vote becomes less attractive for Italy and finding an agreement gets more unlikely due to the unreliability of others.

Under majority rule, the *fear of being excluded* is predominant. This reduces the responders' incentives to act "tough" compared to unanimity rule (see Buchanan and Tullock (1965)). The knowledge that after Italy voting No, France will build a coalition with Germany, makes Italy very "cheap". Though, given that Germany has an established reputation as an unreliable partner, such *fear of being excluded* shrinks. That is, Germany's unreliability makes Italy, similarly as under unanimity rule (though the underlying reasoning differs), less willing to vote yes and an agreement harder to achieve. Consequently, under both voting rules, *a party's unreliability makes others more reluctant to vote yes and thus more "expensive".*

In most cases of interest, take for instance the EU Council or EU Parliament, the bargaining parties sit together and *simultaneously* vote about a proposed issue. There, parties do not observe others' votes before going to the ballot themselves but have formed beliefs about the others' voting intentions. These beliefs are based on the proposal and what has happened before the day of the vote, when the actual bargaining takes place. Parties talk to each other, try to form coalitions, and secure yes-votes. This, I assume to happen *sequentially.* That is, the Proposer (France) approaches the responders (Germany and Italy) one after another, while previous approaches are disclosed. So when facing an offer, responders announce their voting

instead of unanimity rule, more often (see for instance Schuette (2019)).

intention publicly. Hence the responder that is approached next is aware of all the responders that have been approached before, of the offers they have received, and of their announced voting decisions. The binding vote will then take place at the end of a period. In case all parties are reliable, they therefore vote exactly as they have announced. So, unreliability means that a party announces a yes-vote during the bargaining phase but then, with a positive probability, still ends up voting No.

In a world where fully reliable bargaining partners seem to become increasingly rare, understanding the bargaining process with unreliable players is essential. In the first part of the paper, I discuss the differences of unanimous and non-unanimous voting under private information with reliable players. Here, I will especially focus on players' signaling incentives which play an essential role in the underlying reasoning of the impact of unreliability. In a second step, I then try to shed some light on the consequences of others being unreliable under the two voting rules.

I start by reviewing the existing literature. The model is introduced in more detail in Section 3. The subsequent sections proceed to address the model assuming that responders are reliable (Section 4) and unreliable (Section 5), respectively. Section 6 provides a review of the primary findings, followed by some further insights and extensions discussed in Section 7. Finally, Section 8 concludes.

2. Literature

This paper aligns with the broader literature on q-majority rules in multilateral bargaining games with complete information. Specifically, it intersects with models introducing heterogeneity regarding individual players' willingness to agree. For instance, in the classic Baron-Ferejohn bargaining game (Baron and Ferejohn (1989)), complete information is paired with varying discount factors among players. A higher discount factor implies a larger "price" for a player's support. Whereas being expensive is a good thing under unanimity voting, Miller et al. (2018) show that it might be a disadvantage under majority voting, as expensive players risk exclusion from winning coalitions. These findings suggest that when heterogeneity stems from private information, players may prefer to signal toughness (e.g., through lower valuations for having an agreement) under unanimity rule, but not necessarily under non-unanimous voting.

The literature on multilateral bargaining with private information remains limited. Tsai and Yang (2010b) investigate a Baron-Ferejohn game with private information about discount factors, revealing equilibria involving delay and oversized coalitions under majority rule. Piazolo and Vanberg (2024a) examine a model with privately known disagreement values, examining the distinct signaling incentives associated with different voting rules. In their examination of a three-player, two-period game with two different player types, they demonstrate that the ”cheap” type may have an incentive to mimic the behavior of the “expensive” one under unanimity rule. Conversely, under majority rule, the dynamic may shift, with the “expensive” type attempting to appear as a “cheap” one.

Similar patterns emerge in the literature on reputational bargaining. Abreu and Gul (2000) analyze a bilateral bargaining game where players may be “obstinate” types committed to claiming a fixed share of the pie. They illustrate how the presence of such types incentivizes “sophisticated” types to imitate them, i.e., to adopt a tough stance, and thus possibly leading to inefficient delays. Ma (2023) extends this concept to a multilateral majoritarian bargaining game, underscoring how the *threat of exclusion* under majority rule dampens incentives to act tough.

My work builds upon and extends this line of reasoning in two key ways. Firstly, it introduces a model that allows for a focused examination of the behavior of individual players, thereby enabling a detailed discussion of their signaling incentives, confirming the results of Piazolo and Vanberg (2024a).[2] Secondly, it seeks to illuminate the implications of bargaining parties becoming increasingly difficult to rely on. Lewicki and Polin (2013) discuss the general role of trust in negotiations and differentiates between short and long-term consequences of having established a reputation of unreliability. In the short term, deceptive tactics may yield benefits (as noted in Schweitzer and Croson (1999)), while in the long term, opponents are far less inclined to seek a coalition with an unreliable party. The latter is, for instance, confirmed by Boles et al. (2000). They conduct an experiment on the consequences of deceptive behavior in a proposer-responder framework and find that players express a diminished desire to interact in forthcoming periods with players that were assessed as untrustworthy. To the best of my knowledge, this paper represents the first attempt to theoretically explore the effects of a party’s reputation for unreliability on other parties’ voting behavior within a multilateral bargaining framework with private information.

3. Model

Three players bargain over the distribution of a pie of size one. There is one fixed proposer (P, she) and two responders (R1 and R2, he). The game lasts at most

[2]In comparison to Piazolo and Vanberg (2024a), the model of the present paper brings two main benefits. Firstly, it manages to find the same results in a simplified analysis without having to consider mixed strategies. Secondly, it allows to isolate effects of interest and, most importantly, discusses the consequences of the presence of unreliability.

two periods. Each Period $t \in \{1, 2\}$ consists of two sub-periods[3] and is of the same structure: First, P decides which of the responders to approach first. Then, in sub-period ta ($1a$ or $2a$), P makes a publicly observable and (weakly) positive offer to the responder, she decided to approach first. The respective responder announces (also publicly observable) his voting intention, which can either be a "Yes" (Y) or a "No" (N) vote. In sub-period tb, P makes an offer to the other responder, who subsequently announces his voting intention (Y or N). At the end of a period the decisive ballot takes place. P proposes to all responders who announced a Y-vote the offer she has made in the respective sub-period to the respective responder. Responders who announced a N-vote face a zero-offer. R1 then votes as announced. In contrast, the *unreliable* R2 might, in Period 1, vote N even after having announced a Y-vote, resulting in the following definition.

Definition 1. *A party is unreliable if there is a probability $p \in \left(0, \frac{1}{2}\right)$ with that it, in Period 1, votes "No" after having announced a "Yes"-vote. A party is reliable if it always votes as it has announced.*

Note that it is assumed that p is exogenously given. A conceivable extension of the model might treat p as a strategic choice.[4] However, our main interest lies in the effect of the (level of) unreliability of R2 on R1's voting behavior.

P's proposal at the end of Period $t \in \{1, 2\}$ is of the form $x^t = (x_1^t, x_2^t)$. There is an agreement if, in the decisive ballot, at least q responders vote Y. R1 then obtains his offered share x_1^t in addition to an exogenously given valuation of agreement v. P gets the remainder of the pie $1 - x_1^t - x_2^t$. R2, in contrast, is assumed to be of a non-strategic "robot" type that announces a Y-vote in Period t if and only if he is offered at least some fixed and publicly known share $c \in (0, 1)$. If there is no agreement in Period 1, there is an immediate breakdown with probability $1-\delta$. With the continuation probability δ we go into a second round of bargaining, where both responders are reliable. If there is no agreement in Period 2, there is an immediate breakdown with probability 1. In case of a breakdown, all three players have a payoff of zero. Beliefs are formed via Bayes' Rule. See Table 1 for a summary of the bargaining procedure.

[3]Here, the actual "bargaining" takes place. Parties discuss and work on the law that is to be voted on. There, parties can signal/announce their intended voting decision and P discloses the concessions she is willing to make.

[4]In case the level of unreliability depends on the received offer, there might be room for "strategically" making use of being perceived as unreliable. This might be interesting to explore in future research and will be discussed in more detail below.

Period t	Player	Action	Decision process
Before ta	P	Approach Ri first ($i \in \{0, 1\}$)	max utility
ta	P	Offers x_i to Ri	max utility
	Ri	Announce Y or N-vote	max utility (for R1)
tb	P	Offers x_j to Rj	max utility
	Rj	Announce Y or N-vote	Y if at least c (for R2)
Ballot (after tb)	P = *France*	Propose $x^t = (x_1^t, x_2^t)$	$x_k^t = x_k$ if Rk announced Y
	R1 = *Italy*	Vote Y or N	as announced
	R2 = *Germany*	Vote Y or N	switch from Y to N with prob p

Table 1: The Bargaining Process

In the following, I compare majority ($q = 1$) with unanimity voting ($q = 2$). In the course of this, the main focus will lie in the interaction between P and R1. R1's valuation of agreement, denoted as v, is of private information to him. This parameter follows a uniform distribution on the interval $[-1, 1]$, as known to all players. Agreement is welfare-maximizing for certain combinations of v and c (specifically, when $1 + v - c > 0$), whereas breakdown is welfare-optimal for other combinations (when $1 + v - c < 0$).

Equilibrium concept. In this game, R1's strategy must specify his voting announcement for each of his agreement valuations, referred to as types, for every potential proposal and history. A history is considered empty during sub-period $1a$. In $1b$, it includes the offer and the announced vote from sub-period $1a$. During sub-period $2a$ (and $2b$), it encompasses all preceding offers and announced votes, along with the responders' votes in the decisive ballot at the end of Period 1. The strategy of the non-strategic type R2 is predetermined (as detailed earlier) and thus can be disregarded. P's strategy entails determining the responder she approaches first in Period 1 and, for each possible historical sequence, formulating Period 1 and Period 2 proposals, as well as determining the responder she approaches first in Period 2. An equilibrium strategy necessitates maximizing the respective player's expected payoff at every point of the game (subgame perfection). It's worth noting that R1's expected utility from voting Y increases continuously with his valuation of finding an agreement v. Consequently, R1's strategy can be simplified: In each Period, he will opt for a Y-vote on a given offer whenever his valuation of agreement exceeds a certain *cutoff*.[5]

[5]The fundamental difference to the model of Piazolo and Vanberg (2024a) is that in their model, there are only two types of responders: Low types mimic high types' behavior, which is especially beneficial if only a few low types do so. Consequently, there is a range of offers where low types

Cutoff type. Hence, whenever R1 is approached in Period 1, there will be a cutoff value $\hat{v}$ such that R1 votes Y if $v \geq \hat{v}$ and N if $v < \hat{v}$. The type of R1 with a valuation of agreement of $v = \hat{v}$ will be called the *cutoff type*. Such a type is indifferent between voting Y and N. There, the probability of $v = \hat{v}$ is zero, i.e., we can, w.l.o.g., assume that such a responder votes Y. The value of $\hat{v}$ will depend on the offer $x^1 = (x_1^1, x_2^1)$, where $x_1^1 = x$ and $x_2^1 \in \{0, c\}$ as well as on the sequence with that responders are approached.[6] The voting Y probability of R1 from P's perspective is $\frac{\hat{v}+1}{2}$.[7]

Employing backward induction, I initially examine the behavior in Period 2. There, R1 opts to vote Y on an offer x if it yields "immediate" positive utility, i.e., when $x + v \geq 0$. P's expected utility depends on the belief concerning R1's valuation of agreement and thus on the cutoff type. After a Period 1 N-vote of R1, P believes R1's valuation of agreement to be uniformly distributed on $[-1, \hat{v}]$. If R1 voted Y in Period 1, then v is believed to be uniformly distributed on $[\hat{v}, 1]$. So by voting N in Period 1, a responder can signal that he is likely to have a relatively low valuation of finding an agreement and consequently it requires a large Period 2 compensation in order for him to vote Y.[8] In Period 1, analyzing R1's voting behavior requires a more nuanced approach, contingent upon the sequence of responders, R2's expected behavior in Period 1 (including reliability), the continuation probability, and the specifics of the Period 1 offer. The cutoff type function $\hat{v}(x^1)$ captures most of these dependencies. P will consider this function when maximizing her expected utility over x^1, guiding us toward the equilibrium proposal.

In the subsequent analysis, one main focus is on determining the cutoff types under the different setups. The equilibrium cutoff not only gives us the equilibrium proposal but also indicates R1's willingness to vote Y (thus providing us with a "price" for R1) and shows how this willingness is affected by certain parameters, such as the Period 1 offer or the reliability of R2. Consequently, the cutoff type serves as

---

would want to mimic high ones if P expects a low type not to mimic, and low types do not want to mimic if P expects them to mimic. Hence, no pure-strategy equilibrium exists. In a mixed-strategy equilibrium, the proportion of low types that mimic, and thus the signaling value, can be adjusted depending on the offer. Conversely, in the model of the underlying paper, the cutoff type allows for the precise determination of the proportion of the responder's "types" that vote Y, eliminating the need for mixing.

[6]For simplicity, we drop the cutoff type's dependencies in the notation. We will see that on the equilibrium path, the cutoff type will essentially just depend on the offer x that R1 faces in Period 1.

[7]The voting Y probability of R2 is p if $x_2^1 \geq c$ and zero otherwise.

[8]Note that such signaling only works if R1 expects his vote to be pivotal with a positive probability. It turns out that this is indeed true in all cases of interest.

the key piece of information P requires when devising her optimal proposal. Simultaneously, providing the cutoff type offers us a comprehensive equilibrium strategy for R1. This is because the cutoff type perfectly describes R1's behavior in Period 1, while R1's actions in Period 2 follow a consistent pattern: Vote Y on an offer x whenever $x + v \geq 0$, and N otherwise.

Responder's price. R1's Period 1 cutoff type is thereby perfectly correlated with his "price". A higher cutoff type indicates a greater likelihood of R1 voting N, thus rendering him expensive, or, in other words, leading to a higher "price". To avoid any confusion when discussing R1's "price" or whether R1 is considered cheap or expensive, the following definition is provided.

Definition 2. *R1's Period 1 price $z(v)$ is defined as the lowest Period 1 offer that R1 is willing to vote Y on.*

Examining R1's price is particularly insightful when comparing different situations of interest, as it allows us to assess the impact of various parameters on R1's willingness to vote Y. Of particular interest is the effect of R2's unreliability on R1's price under the two voting rules. The following sections thus explore the consequences of the presence of an unreliable party under both majority and unanimity rule. To accomplish this, it is crucial to first understand the signaling incentives under both voting rules when parties are reliable (see Section 4). Subsequently, we are prepared to assess how unreliability impacts these signaling incentives (see Section 5).

4. Reliable parties ($p = 0$)

This section assumes that R2's voting announcement can be fully trusted, i.e., $p = 0$. More specifically, R2 ends up voting Y whenever he is offered at least c (by assumption) and votes N, otherwise.

4.1. Majority rule

First, let us consider majority voting. Throughout both periods, P's offer to R2 will either be zero or c. When P offers c to R2, an immediate agreement is guaranteed since R2 is fully reliable. P strategically leverages the potential for an immediate and certain agreement to heighten R1's *fear of exclusion*. The *threat of exclusion* is most powerful when it looms within the same period,[9] leading to the following Lemma.

[9] Recall that R2 is of a non-strategic type and *fear of exclusion* cannot make him more eager to vote Y. Hence P never approaches R2 first.

Lemma 1. *Under majority rule, P always approaches R1 first. In Period 2, R1 announces a Yes-vote on any (weakly) positive offer.*

If R1 votes[10] N in the sub-period $2a$, P forms a Minimum Winning Coalition (MWC)[11] with R2 in sub-period $2b$. R1 is aware that, independent of his voting decision, there will be an agreement. Consequently, P, in sub-period $2a$, will make a zero-offer to R1 and R1 announces a Y-vote,[12] leading to a certain agreement.

Period 1 behavior is shaped by players' anticipations for Period 2 outcomes. P expects securing an agreement that allows her to retain the entire pie. Consequently, in sub-period $1b$, P faces a trade-off: she must weigh the option of an immediate agreement, costing her c, against the hope for Period 2 to be reached, which happens with probability δ. The former is advantageous if R2 is cheap, while the latter is preferable if the likelihood of advancing to Period 2 is high.

The resolution of this trade-off dictates R1's behavior in sub-period $1a$. So, in the following, let us differentiate between the following two cases.

Firstly, suppose that $c > 1-\delta$. Then P makes a zero offer to R2 ("waits"). R1 anticipates this and therefore does not need to fear immediate exclusion. Consequently, a responder with a negative valuation of agreement can hope for a breakdown to occur (with probability $1-\delta$). More specifically, R1 is indifferent between voting Y and N if $x + v = (1-\delta)0 + \delta v$, i.e., if $v = \hat{v}$ with

$$\hat{v} = -\frac{x}{1-\delta}, \text{ given } c > 1-\delta.$$

Therefore, R1's cutoff type $\hat{v}(x)$ is negative. Thus, a responder with a positive valuation of agreement votes Y on all positive offers x (remember that R1 votes Y whenever $v \geq \hat{v}(x)$). Also, the cutoff value is decreasing in δ, i.e., R1 is more likely to vote Y the larger is the continuation probability. This trend stems from R1's apprehension about being forced to vote Y on a zero offer (similar to being excluded from a coalition), a concern that intensifies with δ.[13]

[10]For R1, and also R2 in case he is reliable, I can use the terms vote and announce a vote likewise. That is because the vote equals the announced vote.

[11]A MWC is characterized by a voting outcome where exactly the necessary number of parties for an agreement to be reached vote Y. Specifically, under majority rule, one responder votes Y while the other votes N.

[12]W.l.o.g., it is assumed that responders vote/announce Y whenever being indifferent between a Y-vote and a N-vote.

[13]Here, P will choose x such that her expected utility $\frac{1+\frac{x}{1-\delta}}{2}(1-x) + \frac{1-\frac{x}{1-\delta}}{2}\delta$ is maximized. This leads to a negative offer (not feasible) and thus P chooses $x = 0$ in the optimum.

Secondly, suppose $c \leq 1 - \delta$. Then, P opts to secure R2's agreement in subperiod $1b$. R1 thus faces a *fear of immediate exclusion* after announcing a N-vote. Analogous to Period 2, this prompts R1 to vote Y on every offer he faces, i.e.,

$$\hat{v} < -1, \text{ given } c \leq 1 - \delta.$$

Combining this with $\hat{v} < -x$ from above (i.e., for $c > 1 - \delta$) leads to the following observation.

Observation 1. *Under majority rule and with reliable responders, R1 fears exclusion and is thus willing to vote Yes on offers that give him negative immediate utility:* $\hat{v} < -x$*.*

The ensuing proposition presents an overview of the resulting equilibrium behavior and follows without further proof.

Proposition 1. *Under majority rule and with reliable responders, P approaches R1 first in both periods and proposes*

- *to R1 in Period* $1a$
  - *if* $c \leq 1 - \delta$*:* $x = 0$*. R1 announces a Yes-vote.*
  - *if* $c > 1 - \delta$*:* $x = 0$*. R1 announces a Yes-vote if* $v \geq 0$*.*
- *to R2 in Period* $1b$*, after R1 announced a No-vote,*
  - *if* $c \leq 1 - \delta$*:* $x_2^1 = c$*. R2 announces a Yes-vote.*
  - *if* $c > 1 - \delta$*:* $x_2^1 = 0$*. R2 announces a No-vote.*
- *to R1 in Period* $2a$*:* $x_1^2 = 0$*. R1 announces a Yes-vote.*
- *to R2 in Period* $2b$*, after R1 announced a No-vote:* $x_2^2 = c$*. R2 announces a Yes-vote.*

Hence P profits substantially from the given sequential structure of the game with reliable responders and majority rule being applied.[14] She benefits most if δ and/or c is small. P benefiting from a low probability of advancing into a second period seems counterintuitive at first. However, a low δ intensifies R1's *fear of exclusion*

[14] In a model with simultaneous voting, there would not be the (same) threat of being "immediately" excluded.

because R1 must expect P to form a MWC with R2 in Period 1 already (the urgency of an immediate agreement for P decreases in δ). So announcing a N-vote becomes unattractive and thus R1 is very "cheap" (his Y-vote essentially comes for "free").[15]

4.2. Unanimity rule

The main feature of unanimity voting is that no party must fear exclusion. Thus, R2 must be offered at least c for an agreement to be feasible. As a result, the sequence in which responders are approached is irrelevant because it is common knowledge that R2 is offered c and announces a Y-vote. Lemma 2 follows directly.

Lemma 2. *Under unanimity rule, P offers c to R2 and is indifferent between approaching either responder first in both periods.*

Period 2 will be reached only after R1 has voted N. Assuming that R1 follows a cutoff strategy, P expects R1's valuation of agreement to be uniformly distributed on $[-1, \hat{v}]$. P's optimal Period 2 proposal depends on R1's cutoff value and is given by

$$x_1^2 = \frac{1 - c - \hat{v}}{2}.$$

Note that x_1^2 is continuously decreasing in $\hat{v}$. The intuition is that the more types of R1 vote Y (smaller $\hat{v}$), the stronger is the signaling value from voting N and thus the larger is the offer made in Period 2. In Period 1, given an offer x, R1 is indifferent between voting Y and N if $v = \hat{v}$, with the cutoff value

$$\hat{v} = \frac{\delta(1-c) - 2x}{2 - \delta}.$$

Consequently, a Period 1 N-vote leads to an increased offer in Period 2: The Period 2 offer amounts $x_1^2 = \frac{(1-\delta)(1-c)+x}{2-\delta}$ (when plugging in $\hat{v}$), which is increasing in x and also larger than x.[16] As a result, R1 is willing to vote N on Period 1 offers that seem profitable at first sight. More specifically, the cutoff type's immediate utility from acceptance would be $x + \hat{v} = \frac{\delta(1-c-x)}{2-\delta}$ and thus is always positive.[17] Hence all types of R1 with $v \in [0, \hat{v}]$ vote N in Period 1 even though they value agreement positively, as stated by the following observation.

[15] Part of this reasoning comes from the assumption that breakdown cannot occur in between offers. A related and also reasonable model would have a breakdown probability following each sub-period, leading to a reduced *fear of "immediate" exclusion* for the responders. Interestingly, this would introduce a cost of forming the "wrong" coalition.

[16] That is because $x + c < 1$ must hold.

[17] $1 - c - x \geq 0$ holds because R2 is offered c and $x + c \leq 1$ must hold.

Observation 2. *Under unanimity rule and with reliable responders, R1 faces, in Period* 1*, a positive signaling value from voting No and thus might vote No on offers that give him positive immediate utility:* $\hat{v} > -x$.

Consequently, Period 1 allows R1 to make the following gamble: Vote N, thereby signal that he must be compensated heavily for a Y-vote (i.e., that he has a low valuation of agreement), and hope to receive a larger share in Period 2. There, the respective *positive signaling value from voting N* increases with δ, rendering R1 more inclined to vote N (i.e., more costly) in Period 1 (as can be seen by $\frac{\partial \hat{v}}{\partial \delta} > 0$). That is because responders can only benefit from signaling if a second round of bargaining is reached.

In a lest step, P's optimal Period 1 proposal is to be computed. This is presented in the proof of the following proposition which summarizes the equilibrium behavior.

Proposition 2. *Under unanimity rule and with reliable responders, P is indifferent between approaching either responder first in both periods and offers c to R2. To R1, she proposes*

- *in Period* 1*:* $x = \max\{0, \tilde{x}(c, \delta)\}$*, where* $\tilde{x}(c, \delta)$ *is decreasing in c. R1 votes Yes if he values agreement sufficiently* ($v \geq \hat{v}$)*.*
- *in Period* 2*, after R1 has voted No:* $x_1^2 > x$*, where* x_1^2 *is increasing in the Period* 1 *offer* x.[18] *R1 votes Yes iff* $x_1^2 + v \geq 0$.

Proof. See the Appendix for the proof of this Proposition and also for the above discussed derivations of x_1^2 and $\hat{v}$. □

Note that the impact of the continuation probability δ on the optimal Period 1 proposal is ambiguous. That is because a reduced Period 1 propensity of R1 to vote Y, which can be caused by a larger δ, induces opposing effects. On one hand, P seeks to take countermeasures against the reduced likelihood of R1 to vote Y and thereby must augment her offer. On the other hand, each type of R1, due to the increased *positive signaling value from voting N*, demands a larger offer. This diminishes the incremental value of an increased offer and potentially leads to a reduction in the optimal offer.

[18] The Period 1 offer is $\tilde{x}(c, \delta) = \frac{\delta(1-\delta)-(2-\delta-\frac{\delta^2}{2})c}{4-3\delta}$. The increased Period 2 offer, after R1 has voted N in Period 1, is increasing in the Period 1 offer x and amounts to $x_1^2(x) = \frac{(1-\delta)(1-c)+x}{2-\delta}$.

5. Unreliable parties ($p > 0$)

In this section, R2 is assumed to be unreliable, i.e., $p > 0$. In Period 1, R2 might now vote N even after having announced a Y-vote (see Definition 1). In the following, the equilibrium behavior under both voting rules and with an unreliable R2 will be analyzed.

5.1. Majority rule

Recall that R2 is only unreliable in Period 1. Hence, in Period 2, nothing has changed compared to the analysis in Section 4.1. Specifically, Lemma 1 remains valid: R1 is approached first in both periods, and in Period 2, R1 votes Y on any offer. Consequently, P retains the entire pie for herself in Period 2. In sub-period 1b, P again - like in the "reliable" case - faces the trade-off of either seeking an immediate agreement with R2 or deferring negotiations to Period 2.

The resolution of this trade-off will be again crucial for R1's Period 1 behavior. Therefore, we must differentiate between the following two cases, exactly like in the "reliable" case above.

Firstly, suppose $c > 1-\delta$, i.e., that R2 is quite costly. As a consequence, P makes a zero-offer to R2 in sub-period 1b. In sub-period 1a, R1 anticipates this. Hence there is no *fear of immediate exclusion* and R1's cutoff type is the same as in the reliable case

$$\hat{v} = -\frac{x}{(1-\delta)}, \text{ given } c > 1-\delta.$$

Secondly, suppose that $c \leq 1-\delta$. Then, P tries to reach an immediate agreement in sub-period 1b. Consequently, P proposes c to R2, who then announces a Y-vote. However, in contrast to the reliable case, there is a chance that R2 might still vote N (with probability p). Hence, in sub-period 1a, R1's concern about a Period 1 exclusion, and thus his cutoff type, hinges on R2's unreliability p:

$$\hat{v} = -\frac{x}{(1-\delta)p}, \text{ given } c \leq 1-\delta.$$

If $x \geq (1-\delta)p$, then $\hat{v} \leq -1$. In words, this means that if the offer is large enough, all types of R1 vote Y in Period 1. As discussed before, the same holds in case R2 is very reliable: Recall that $p = 0$ leads to $\hat{v} < -1$. Increasing R2's unreliability p leads to a reduced threat of Period 1 exclusion for R1. Thus, R1 becomes more willing to vote N.[19] The same we can observe by looking at R1's Period 1 price:

[19] Also note that if δ is large, breakdown becomes unlikely, and thus R1 is again willing to vote

Observation 3. *Suppose majority rule is applied, $c \leq 1-\delta$, R2 is unreliable, and R1 values agreement negatively $v < 0$.*[20] *Then R1's Period* 1 *price $z(v) = -(1-\delta)pv$ is increasing in R2's level of unreliability p.*

Recall that we defined the price of R1 with a valuation of agreement v as the offer, where there is a continuation equilibrium with R1 being indifferent between voting Y and N in Period 1 (see Definition 2). So, discussing R1's cutoff value is equivalent to discussing his price. Note that the above observation only holds if R2 is not too expensive and $v < 0$. In the remaining cases, the relationship between R2's unreliability and R1's price is the following: If R2 is expensive, his price will be independent of p. If R1 values agreement positively and R2 is cheap, then R1 votes Y on all offers, i.e., his "price" is zero and thus also independent of p.

Solving P's maximization problems in both cases (for small c and large c, respectively) leads to Proposition 3.

Proposition 3. *Under majority rule and with an unreliable R2, P approaches R1 first in both periods and proposes*

- *to R1 in Period* 1*a*
  - *if $c \leq 1-\delta$: $x = \tilde{x}(c,\delta,p) > 0$.*[21] *R1 votes Yes if $v \geq = -\frac{\tilde{x}(c,\delta,p)}{(1-\delta)p}$.*
  - *if $c > 1-\delta$: $x = 0$. R1 votes Yes if $v \geq 0$.*
- *to R2 in Period* 1*b, after R1 announced a No-vote,*
  - *if $c \leq 1-\delta$: $x_2^1 = c$. R2 announces a Yes-vote.*
  - *if $c > 1-\delta$: $x_2^1 = 0$. R2 announces a No-vote.*
- *to R1 in Period* 2*a: $x_1^2 = 0$. R1 announces a Yes-vote.*
- *to R2 in Period* 2*b, after R1 announced a No-vote: $x_2^2 = c$. R2 announces a Yes-vote.*

Proof. See Appendix. □

Y on small offers. That is because reaching Period 2 is bad for R1 as then he will have to vote Y on the smallest possible offer (zero).

[20] Note that R1's price is of our main interest if he values agreement negatively. Otherwise, he will just vote Y on any offer.

[21] $\tilde{x}(c,\delta,p) = \min\{\frac{(1-p)c}{2}, (1-\delta)p\}$.

If R1 values agreement positively, he votes Y in Period 1 with certainty, resulting in an immediate agreement. If R1 values agreement negatively, his Period 1 behavior hinges on his *fear of exclusion* or, more specifically, on his hope for not finding an agreement (i.e., a breakdown). As the likelihood of a breakdown increases, R1's willingness to vote Y decreases. Thus, R2's unreliability p, similar to the breakdown probability $1 - \delta$, raises R1's price and may prompt a positive Period 1 offer.

5.2. Unanimity rule

If unanimity rule is applied, P can only secure an agreement if she offers c to R2. R2 will then announce a Y-vote but still ends up voting N with probability p. This outcome remains consistent regardless of the order in which responders are approached, meaning Lemma 2 remains valid.

Moreover, the process driving the players' equilibrium behavior remains largely analogous to the reliable case discussed above. In Period 1, R1 once again confronts a trade-off: either opt for a N-vote in the hope for an increased Period 2 offer, or opt for a Y-vote, thereby pursuing immediate agreement. The pivotal distinction now is that a Y-vote no longer ensures immediate agreement with certainty, due to R2's unreliability. As a consequence, we must not only discuss the Period 2 proposal after R1 has voted N but also after R1 has voted Y. After such a Y-vote, P believes R1's valuation of agreement to lie in the interval of $[\hat{v}, 1]$. She will then end up making a zero offer if Period 2 is reached (see proof of Proposition 4 for more details). The Period 2 offers are therefore given by

$$x_1^2 = \begin{cases} \frac{1-c-\hat{v}}{2} & \text{after Period 1 N-vote} \\ 0 & \text{after Period 1 Y-vote.} \end{cases}$$

Thus, P and R1 must not only take into account the signaling value from voting N but also the one from voting Y. Here, a prior Y-vote signals to P that R1 is likely to value agreement positively, leading to a smaller Period 2 offer than a Period 1 N-vote and thus to the following Observation.

Observation 4. *Suppose unanimity rule is applied and R2 is unreliable. Then, in Period* 1*, R1 faces a positive signaling value from voting No and a negative signaling value from voting Yes.*

Thus, a Y-vote carries the risk of receiving a low offer in a subsequent period. Consequently, R2's unreliability decreases R1's willingness to vote Y, resulting in an increased Period 1 price - as described by the following Observation:

Observation 5. *Suppose unanimity rule is applied, R2 is unreliable, and* $v \in [c-1, 1-c]$. *Then R1's Period* 1 *price* $z(v) = -(1-\delta)v + \frac{\delta(1-c-v)}{2(1-p)}$ *is increasing in R2's level of unreliability* p.[22]

Proof. See Lemma 7 for the derivation of R1's price. $\frac{\partial z(v)}{\partial p} = \frac{(1-c-v)\delta}{2(1-p)^2}$ is positive if $v < 1-c$. If $v > 1-c$, R1 votes Y with certainty (price of zero). If $v < c-1$, R1 votes N with certainty (price larger than one). □

Recall that R1's Period 1 price comes from the same constraint as his cutoff type. A higher cutoff type implies that fewer types of R1 are willing to vote Y in Period 1, resulting in a higher Period 1 price. In this case, the equilibrium cutoff type is given by[23]

$$\hat{v} = \frac{\delta(1-c) - 2(1-p)x}{2(1-p)(1-\delta) + \delta}.$$

Similar to R1's price, this cutoff value is increasing with R2's unreliability p.

Proposition 4 now outlines the equilibrium behavior, derived from solving P's maximization problem(s) while taking R1's respective cutoff type into account.

Proposition 4. *Under unanimity rule and with an unreliable R2, P is indifferent between approaching either responder first in both periods and offers c to R2. To R1, she proposes*

- *in Period* 1*:* $x = \max\{0, \tilde{x}(c,\delta,p)\}$, *where* $\tilde{x}(c,\delta,p)$ *is decreasing in c. R1 votes Yes if* $v \geq \hat{v}$.
- *in Period* 2*, after R1 announced a Yes-vote in Period* 1*:* $x_1^2 = 0$. *R1 votes Yes if* $v \geq 0$.
- *in Period* 2*, after R1 announced a No-vote in Period* 1*:* $x_1^2 > x$.[24] *R1 votes Yes if* $x_1^2 + v \geq 0$.

[22]The here discussed price $z(v)$ is the on-equilibrium path price, where $\hat{v} \geq 0$. If $\hat{v} < 0$, then R1's price would amount $z(v) = -v + \frac{\delta(1-c+v)}{2(1-p)}$. However, making an offer that leads to $\hat{v} < 0$ is dominated for P. For more details see the Appendix or the next footnote.

[23]Note that $\hat{v} = \frac{\delta(1-c)-2(1-p)x}{2(1-p)-\delta}$ if $\hat{v} < 0$. However, P's equilibrium proposal is only consistent with $\hat{v} \geq 0$ (in other words: there is no equilibrium with $\hat{v} < 0$), where $\hat{v} = \frac{\delta(1-c)-2(1-p)x}{2(1-p)(1-\delta)+\delta}$. So the latter $\hat{v}$ is the one we concentrate on.

[24]The optimal Period 1 offer is $x = \max\{0, \tilde{x}(c,\delta,p)\}$, where $\tilde{x}(c,\delta,p) = \frac{(1-\delta)\delta-(1-\delta)c(2+\delta-2p+2p\delta)-\frac{\delta^2}{2(1-p)}c}{4(1-p)(1-\delta)+\delta}$. The optimal Period 2 offer, after R1 has voted N in Period 1, is $x_1^2 = \frac{1-c-\hat{v}}{2}$.

Proof. See Appendix. □

As noted above, the effect of R2's unreliability on R1's price is clearly identified and positive (see Observation 5). In contrast, the impact on P's Period 1 offer is ambiguous. This ambiguity arises because a higher price for R1 does not necessarily result in a larger offer from P, although it might. Further insights on the effect of R2's unreliability are discussed in the subsequent section.

6. A comparison of the voting rules: Main results

In the following, I discuss the differences between the two voting rules, particularly focusing on the distinct strategic considerations R1 takes into account. Additionally, I (re-)state the primary findings of the previous analyses. Under unanimity rule, the emphasis is on understanding R1's behavior for intermediate values of v. If R1 highly values reaching an agreement, he will vote Y regardless of the risks, as the potential breakdown is not worth the gamble. Conversely, if R1 strongly opposes the agreement, he will consistently vote N, knowing that no agreement will materialize without his consent. Although the following results apply to all values of v, our focus when discussing the main results narrows to the more insightful range $v \in (c-1, 1-c)$.[25]

First, let us compare the responders' inclination to vote Y (i.e., R1's prices) and the associated likelihood of breakdown between majority and unanimity rule. The disparities primarily stem from the *fear of exclusion*[26] under majority rule and the *positive signaling value from voting N* under unanimity rule. These factors, influenced in part by private information,[27] exert opposing effects on R1's willingness to vote Y, i.e., on his Period 1 price: Under majority rule, R1's price is $z(v) = -(1-\delta)pv$ if c is offered to R2 in sub-period $1b$ (i.e., if $c \leq 1-\delta$) and $z(v) = -(1-\delta)v$ if R2 is offered zero in $1b$ (i.e., if $c > 1-\delta$). Under unanimity rule, R1's price is $z(v) = -(1-\delta)v + \frac{\delta(1-c-v)}{2(1-p)}$. This leads to the following two results (for more details, see Appendix B).

[25]Note that $v > c-1$ means that having an agreement is efficient. When discussing welfare, we will also take a look at $v < c-1$.

[26]To be more precise, in our setup the *fear of exclusion* is the fear of ending up in an agreement which results in a negative utility (i.e., where $x + v < 0$). This can either be the case when R1 is excluded from the coalition or when R1 is pressured to vote Y on small offers that lead to such negative utility. The latter is always the case when Period 2 is reached under majority voting.

[27]There is no room for any signaling under complete information. Consequently, the differences between the two voting rules are larger in the private information environment than in the full information one.

Result 1. *When facing a Period 1 offer x, R1 is (weakly) more willing to vote Yes if majority rule is applied than if unanimity rule is applied.*

Finding an agreement under unanimity rule is - by construction - harder to achieve than under majority rule: By definition more players must be convinced to agree if unanimity is required. In addition, each responder individually is less willing to vote Y under unanimity than under majority rule (see Result 1). Thus, Result 2 is not surprising.

Result 2. *The probability of breakdown is lower under majority rule than under unanimity rule.*

It's noteworthy that if R2 is fully reliable and unanimity rule is applied, there is a breakdown in half of the cases independent of the parameter values - hence, even when δ is large and the cost of R2 (c) is low. In contrast, under majority rule, there might be an agreement with certainty in such cases. This outcome is due to R1 becoming prohibitively expensive under unanimity rule. When R1 perceives a diminished risk of breakdown (with a large δ), he is more inclined to take the risk to vote N in Period 1 with hopes of securing a larger share in Period 2 - a scenario characterized by a substantial *positive signaling value from voting N*. Similarly, if R2's cost is low, R1 can anticipate a larger share in Period 2, incentivizing him to vote N in Period 1. At the same time, he faces a larger Period 1 offer, ultimately leading to a stable probability of agreement.

While the results 1 and 2 hold independent of R2's reliability, it is crucial to examine the effects of R2's level of (un)reliability on the responders' and the Proposer's behavior. The following pivotal finding of our analysis results directly from the observations 3 and 5.[28]

Result 3. *The level of R2's unreliability p (weakly) increases R1's price under both voting rules.*

The reasoning underlying this result differs between the two voting rules. Under unanimity rule, the unreliability of one responder makes a Y-vote less attractive for the other responder: Instead of leading to a certain immediate agreement, a Y-vote now signals that R1 values agreement. So a Period 1 Y-vote leads to a lower Period 2 offer than a Period 1 N-vote, thereby reducing the attractiveness of voting Y. Hence, R1's Period 1 price increases because of this *negative signaling value from*

[28] In the cases that are not included in these two observations R2's unreliability p does not affect R1's price.

voting Y. Under majority rule, the unreliability of R2 makes reaching an agreement in sub-period 1*b* less likely, thereby reducing R1's *fear of immediate exclusion*. This decreases R1's "costs" of a N-vote in Period 1*a* and thus leads to a larger Period 1 price of R1.

As a consequence, the presence of an unreliable party complicates the process of reaching agreements. This is attributed not only to the reduced probability of a Y-vote from the unreliable party but also to the inflated prices demanded by the other responder. However, the impact on the equilibrium Period 1 offers is less evident because an increased price does not necessarily lead to an increased offer. Under majority rule, P has increased incentives to convince R1 to vote Y. P does not want to solely rely on the unreliable R2 and thus proposes a (weakly) larger offer to R1 (see Proposition 3). At the same time, P's increased willingness to form a coalition with R1, reduces R1's *fear of being excluded* and thus makes him more expensive. So, even though P increases the offer when p becomes larger, R1 is less likely to vote Y in Period 1, leading to more inefficient breakdowns. Conversely, under unanimity rule, the impact of R2's unreliability on the Period 1 offer is ambiguous (see Proposition 4), whereas the probability of breakdown unambiguously increases. All of these insights are outlined by the subsequent result.

Result 4. *Under majority rule, the unreliability of R2 (weakly) increases the Period* 1 *equilibrium offer to R1 and decreases the overall probability of finding an agreement.*

Under unanimity rule, the unreliability of R2 has an ambiguous effect on the Period 1 *equilibrium offer to R1 and decreases the overall probability of finding an agreement.*

Thus, the unreliability of R2 does not inherently result in a lower proposal to R1. However, it does decrease the likelihood of both parties voting Y for a given offer. Consequently, if finding an agreement is efficient, the unreliability of one responder has "additional" negative implications on the overall welfare because it increases the probability of inefficient breakdowns. These negative implications come in "addition" to the "mechanical" effect of one of the parties being unreliable (and therefore by definition being less likely to vote Y) and are based on the "endogenously" decreased willingness of voting Y of the reliable party.

7. Extensions and further insights

In this section, I explore additional insights that do not directly follow from the presented model but require slight adjustments or extensions to the model. Details and examples are provided in Appendix C.

7.1. Larger coalitions

In the model discussed above, MWCs are the only equilibrium outcome under majority rule, even when one of the responders is unreliable. This is specific to the three-player case. If the rational responder R1 announces a Y-vote, it leads to an immediate agreement, giving P no incentive to make a positive offer to the other responder. Although a general analysis is beyond the scope, I consider a five-player model in Appendix C.

First, consider a simplified version of our original model, where R1 - just like R2 - is an unreliable robot type. In Period 1, P will form an oversized coalition (OSC)[29] and offer c to both responders if

$$c < \frac{p\,(1-\delta(1-c))}{1+2p} \text{ and } \delta < \frac{1-2c}{1-c}.$$

OSCs are thus formed when responders are cheap (small c), unreliable (large p), and the probability of a second chance to propose is low (low δ).

Second, consider a more complex scenario with also a rational responder being part of the model. It can be shown that P might still prefer to form an OSC, with the rational responder being included in the coalition. Suppose there are two additional responders (resulting in five players in total), with two of the four responders (say, R2 and R3) being of the unreliable robot type, just like R2 in the original model. Assume R4 is too expensive to be convinced of a Y-vote and R1 is a "rational" responder, with a valuation of agreement that is continuously distributed on $[-1, 0]$. For simplicity, let δ be zero (i.e., there is only one period).[30] In equilibrium, P forms an OSC with R1, R2, and R3 if

$$c < \frac{p}{1+3p}.$$

Specifically, P offers $x = \frac{p}{1+p}(1-2c)$ to R1 and c to R2 and R3.[31] If p is very small,

[29]An OSC is defined as a coalition where more parties receive a positive offer than the at least required number of parties for an agreement to be achieved. In the three player case, an OSC means that positive offers are made to both responders.

[30]As discussed above, OSCs are good options as long as δ is small, i.e., the assumption $\delta = 0$ is not required. However, I use it to simplify the equilibrium analysis.

[31]Note that R1's price is $z = -\frac{2p}{1+p}v$ and thus increasing in p. Consistent with the original model, the unreliability of the other responder(s) makes R1 more expensive. The optimal proposal being offered to R1 is also increasing in p, i.e., R1 is being compensated for the heightened risk of proposal failure.

then P is better off proposing a MWC. Thus, the unreliability of responders indeed increases the likelihood that P forms an OSC, as noted in the following observation.

Observation 6. *In an extended model with five players, the presence of unreliable responders leads to (weakly) more parties being included in the winning coalition under majority rule than with reliable parties only.*

Including more responders than the minimum required in the coalition reduces the risk of inefficient breakdowns. This strategy is particularly relevant when δ and c are small, i.e., if it is unattractive for P to wait for another round of bargaining and if the unreliable responders are not too expensive. Consequently, in scenarios where parties are not entirely reliable, the occurrence of MWCs should be significantly lower than what theory predicts for situations with fully reliable responders.

The reason why unreliability leads to OSCs is that it leads to uncertainty about the responders' voting behavior. Such uncertainty can also stem from other sources, such as private or incomplete information. As theoretically demonstrated by Tsai and Yang (2010a), uncertainty alone can drive the formation of oversized coalitions under majority rule. Experimental evidence supporting this finding is provided by Piazolo and Vanberg (2024b).

7.2. Asymmetric beliefs about the future

The assumption of incorrect and asymmetric beliefs about δ aligns with real-world scenarios filled with uncertainty, such as negotiations for new climate policies. In these negotiations, different parties may hold varying (and incorrect) expectations regarding the urgency of implementing new measures, influencing their expected probabilities of still reaching an "early enough" new treaty after an initial proposal failure.

The continuation probability δ is a key driver for voting N under unanimity rule. It increases the *positive signaling value from voting N*, making R1 more expensive. Incorrect and asymmetric beliefs about δ can therefore lead to an overestimation of the *positive signaling value from voting N*, possibly resulting in (too) many rejections of Period 1 proposals. In such cases, if parties highly value reaching an agreement, the potential costs of a breakdown under unanimity rule may outweigh the costs of a possible exclusion under majority voting. Consequently, when responders hold asymmetric beliefs about the continuation probability, a shift from unanimity to majority voting may be mutually beneficial for all players, i.e., a Pareto improvement (refer to Lemma 9 in Appendix C).

8. Concluding remarks

In the realm of multilateral bargaining under private information, the bargaining parties' strategic considerations depend on the voting rule that is applied. Under unanimity rule, the act of voting "No" carries a distinct *positive signaling value*, making responders costly and agreements hard to reach. Conversely, under majority rule, the *fear of exclusion* associated with voting "No" makes responders more inclined to vote "Yes", thereby enhancing the likelihood of agreement. Thus, transitioning from unanimity to majority rule consistently favors the proposer and mitigates the risk of inefficient breakdowns and delays. However, such a shift does in general not qualify as a Pareto improvement (with a potential exception being a scenario involving asymmetric beliefs about the continuation probability, as discussed above).

Despite the advantages of non-unanimous voting rules, parties suffer from the *fear of being excluded*, i.e., through the formation of minimum winning coalitions. While theoretical models often anticipate the prevalence of minimum winning coalitions, empirical evidence presents a different narrative. Both laboratory experiments[32] and real-world observations in institutions like the EU Parliament reveal fewer instances of exclusions than predicted. Within the EU Council, where non-unanimous voting is employed, a remarkable "consensus rate" exceeding 70% is achieved (see Hayes-Renshaw et al. (2006)).[33] So even when majority rule is applied, there appears to be a strong inclination among EU member states (as well as lab participants) toward seeking unanimity. In an extended model discussed in the previous section, we theoretically confirm that the presence of unreliable parties may indeed increase the likelihood of an "oversized coalition" to be formed. Unreliability can thus serve as a possible explanation for the deviations between the theoretical predictions of standard bargaining models and empirical observations. However, even though minimum winning coalitions may appear less often than theoretical predicted, the prospect of potential exclusions still serves as a catalyst for increasing parties' willingness to agree and compromise. Thus, while the possibility of facing exclusion remains a drawback, it is counterbalanced by the advantages of majority voting, namely, the increased likelihood of reaching agreements and thereby enhancing an institution's ability to act.

Despite the advantages of non-unanimous voting rules, parties suffer from the fear of being excluded, particularly through the formation of minimum winning coalitions.

[32] Piazolo and Vanberg (2024b) find in a lab experiment that MWCs are proposed in around only one-third of the cases, contrary to theoretical predictions.

[33] These statistics pertain to the period from 1994 to 2004, with recent years witnessing a consensus rate surpassing 80% (refer to Mintel and Ondarza (2022)).

While theoretical models frequently predict the prevalence of minimum winning coalitions, empirical evidence tells a different story. Both laboratory experiments[34] and real-world observations in institutions like the EU Parliament show fewer instances of exclusions than anticipated. In the EU Council, where non-unanimous voting is employed, a remarkable "consensus rate" exceeding 70% has been achieved (see Hayes-Renshaw et al. (2006)).[35] Even when majority rule is applied, there seems to be a strong preference among EU member states (and lab participants) for reaching unanimous agreements. In the extended model discussed earlier, we theoretically demonstrate that the presence of unreliable parties may increase the likelihood of forming an "oversized coalition". Unreliability thus offers a potential explanation for the discrepancy between standard bargaining models' predictions and empirical observations. However, while minimum winning coalitions appear less frequently than predicted, the threat of potential exclusion still serves as a catalyst for increasing parties' willingness to agree and compromise. Thus, while the risk of exclusion remains a drawback, it is counterbalanced by the benefits of majority voting, including the higher likelihood of reaching agreements and strengthening an institution's capacity to act.

This dynamic gains particular relevance in the presence of unreliable parties. In a world where bargaining parties' announced voting decisions cannot be fully relied on, reaching an agreement becomes ever more difficult. This difficulty arises not only from the inherent decrease in the likelihood of parties voting "Yes", but, more importantly, from the main result of this paper: The presence of an unreliable party decreases the other parties' willingness to vote "Yes". Under majority rule, this is based on a reduced *fear of being excluded*. Reaching an agreement with an unreliable party is harder to be achieved and thus another party's exclusion becomes less likely. Consequently, larger offers are required and the formation of more stable coalitions that include additional parties beyond the minimum requirement is incentivized. Under unanimity rule, in contrast, the larger responders' prices are due to the more pronounced *negative signaling value from voting "Yes"*. Independent of the applied voting rule, the presence of an unreliable Germany therefore poses an increased difficulty for France to persuade Italy to vote "Yes".

Future research could, on one hand, explore an adjusted model in which Germany's level of unreliability is endogenously determined, dependent on the offer re-

[34] Piazolo and Vanberg (2024b) find that minimum winning coalitions are proposed in only about one-third of cases, contrary to theoretical expectations.

[35] This statistic refers to the period from 1994 to 2004, with recent years showing a consensus rate of over 80% (see Mintel and Ondarza (2022)).

ceived, and can thus be used “strategically”.[36] This approach would provide further insights into the (un)attractiveness of being perceived as an unreliable bargaining partner, which is likely to vary between different voting rules (see Abreu and Gul (2000) and Ma (2023)).[37] On the other hand, future research could explore related theoretical setups where a share of bargaining partners have non-standard preferences.[38] Such theoretical analyses would benefit greatly from being complemented by empirical and experimental studies investigating the effects of unreliability in bargaining situations under private information.

Acknowledgements

I would like to thank Christoph Vanberg for his valuable support at all stages of this project and seminar participants at the CMID in Budapest for their comments.

[36]Such model would include Germany as a rational player instead of a robot type.

[37]Abreu and Gul (2000) show that being perceived as a tough negotiator is advantageous under unanimity rule. In contrast, Ma (2023) demonstrates that under majority voting, the potential for exclusion makes it costly to be seen as an unattractive bargaining partner.

[38]Cooper and Kagel (2016) discusses other-regarding preferences, noting that our understanding of such dynamics in bargaining games is still incomplete. For example, Montero (2007) shows that in certain conditions under the inequality-aversion model by Fehr and Schmidt (2001), the fear of exclusion may even intensify in bargaining games under majority rule with inequality-averse players.

References

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

Baron, D., Ferejohn, J., 1989. Bargaining in legislatures. American political science review 83, 1181–1206.

Boles, T., Croson, R., Murnighan, K., 2000. Deception and retribution in repeated ultimatum bargaining. Organizational behavior and human decision processes 83, 235–259.

Buchanan, J.M., Tullock, G., 1965. The calculus of consent: Logical foundations of constitutional democracy. volume 100. University of Michigan press.

Cooper, D.J., Kagel, J.H., 2016. Other-regarding preferences. The handbook of experimental economics 2, 217.

Fehr, E., Schmidt, K.M., 2001. Theories of fairness and reciprocity-evidence and economic applications. Available at SSRN 264344 .

Hayes-Renshaw, F., Van Aken, W., Wallace, H., 2006. When and why the EU council of ministers votes explicitly. JCMS: Journal of Common Market Studies 44, 161–194.

Human Rights Watch, 2024. Germany throws a spanner in works of the eu supply chain law. `https://www.hrw.org/news/2024/01/29/germany-throws-spanner-works-eu-supply-chain-law`. Accessed: 2024-03-30.

Lewicki, R., Polin, B., 2013. Trust and negotiation, in: Handbook of research on negotiation. Edward Elgar Publishing, pp. 161–190.

Ma, Z., 2023. Efficiency and surplus distribution in majoritarian reputational bargaining. Journal of Economic Theory 210, 105649.

Miller, L., Montero, M., Vanberg, C., 2018. Legislative bargaining with heterogeneous disagreement values: Theory and experiments. Games and Economic Behavior 107, 60–92.

Mintel, J., Ondarza, N.v., 2022. More EU Decisions by qualified majority voting-but how? Legal and political options for extending qualified majority voting. SWP Comment 61. German Institute for International and Security Affairs. Berlin.

Montero, M., 2007. Inequity aversion may increase inequity. The Economic Journal 117, C192–C204.

Piazolo, D., Vanberg, C., 2024a. Legislative bargaining with private information: A comparison of majority and unanimity rule. Working Paper 753. Heidelberg University. Heidelberg.

Piazolo, D., Vanberg, C., 2024b. Multilateral bargaining with private information – experimental results on the comparison of different voting rules. Unpublished.

Schnedler, W., Vanberg, C., 2014. Playing 'hard to get': An economic rationale for crowding out of intrinsically motivated behavior. European Economic Review 68, 106–115.

Schuette, L., 2019. Should the EU make foreign policy decisions by majority voting. Center for European Reform .

Schweitzer, M., Croson, R., 1999. Curtailing deception: The impact of direct questions on lies and omissions. International Journal of Conflict Management 10, 225–248.

Tsai, T.S., Yang, C., 2010a. Minimum winning versus oversized coalitions in public finance: The role of uncertainty. Social Choice and Welfare 34, 345–361.

Tsai, T.S., Yang, C., 2010b. On majoritarian bargaining with incomplete information. International Economic Review 51, 959–979.

Appendix A. Proofs of main propositions

In the following, the proofs of the Propositions 2, 3, and 4 are presented. Recall that Proposition 1 has followed from the discussions before and has thus been stated without any further proof. The structures of the proofs are very much alike. They use backward induction and present Lemmas representing the results of the following steps. First, Period 2 behavior is analyzed. Then, the Period 1 cutoff type is calculated. Lastly, P's Period 1 maximization problem is solved, leading to the optimal proposals.

Appendix A.1. Proof of Proposition 2

Lemma 3. *In Period* 2*, under unanimity rule, P proposes*

$$x^2 = \begin{cases} (0, c) & \hat{v} > 1 - c \\ (\frac{1-c-\hat{v}}{2}, c) & \hat{v} \in [c-1, 1-c] \\ FAIL & \hat{v} < c - 1. \end{cases}$$

Proof. Given $x^2 = (x_1^2, c)$, R1's probability of voting Y is $\frac{\hat{v}+x_i^2}{\hat{v}+1}$ and P's Period 2 utility thus amounts to $EU_P(x^2 = (x, c)) = \frac{\hat{v}+x}{\hat{v}+1}(1 - x - c)$. Maximizing this over x leads to $x = \frac{1-c-\hat{v}}{2}$. Here $x + c = \frac{1+c-\hat{v}}{2} \leq 1 \Leftrightarrow c - \hat{v} \leq 1$ (P only has a budget of one) and $x \geq 0 \Leftrightarrow c + \hat{v} \leq 1$ (offers must be positive) are required, i.e., $\hat{v} \in [c-1, 1-c]$. If $\hat{v} < c-1$, then P prefers to offer a proposal that fails for sure $x^2 = FAIL$ (agreement is just too expensive) leading to $EU_P(x^2 = FAIL) = 0$. If $\hat{v} > 1 - c$, then P offers $x^2 = (0, c)$ leading to $EU_P(x^2 = (0, c)) = \frac{\hat{v}}{\hat{v}+1}(1 - c)$. Given that $\hat{v} \in [c - 1, 1 - c]$, $EU_P(x^2 = (x, c)) = \frac{(1-c+\hat{v})^2}{4(1+\hat{v})}$. □

In Period 1, P basically has two options. She can either propose an offer that is sure to fail $x^1 = FAIL$ (e.g., by choosing $x_2^1 < c$), which thus leads to $\hat{v} = 1$, or she can offer $x^1 = (x_1^1, c)$.

Lemma 4. *Given $x^1 = (x, c)$, R1 votes Y iff $v \geq \hat{v} = \frac{\delta(1-c)-2x}{2-\delta}$.*

Proof. R1 is indifferent between voting Y and N if

$$x + v = \delta \begin{cases} \max\{v, 0\} & \hat{v} > 1 - c \\ \max\{v + \frac{1-c-\hat{v}}{2}, 0\} & \hat{v} \in [c-1, 1-c] \\ 0 & \hat{v} < c - 1 \end{cases}$$

(LHS is R1's utility from voting Y and RHS from voting N) and $v = \hat{v}$. Hence R1's cutoff type is $\hat{v} = \begin{cases} \frac{\delta(1-c)-2x}{2-\delta} & \hat{v} \in [c-1, 1-c] \\ -\frac{x}{1-\delta} & \hat{v} > 1-c \\ -x & \hat{v} < c-1 \end{cases}$. Note that we can not be in the upper case because $\hat{v} \leq 0$ contradicts $\hat{v} > 1-c$. Also note that $x \in [0, 1-c]$ and we thus can also not be in the lower case. Hence, as $x \in [0, 1-c], \hat{v} = \frac{\delta(1-c)-2x}{2-\delta} \in [c-1, \frac{\delta}{2-\delta}(1-c)]$. □

Lemma 5. *P's optimal Period* 1 *offer is* $x^1 = \left(\max\left\{0, \frac{\delta(1-\delta)-(2-\delta-\frac{\delta^2}{2})c}{4-3\delta}\right\}, c\right)$.

Proof. Recall from above that P's expected Period 2 utility after a N-vote of R1 is $EU_P^2 = \frac{(1-c+\hat{v})^2}{4(1+\hat{v})}$. Also, P knows that $\hat{v} = \frac{\delta(1-c)-2x}{2-\delta}$ and will in Period 1 thus choose x such that $EU_P((x,c),((\frac{(1-\delta)(1-c)+x}{2-\delta}, c))) = \frac{1-\hat{v}}{2}(1-x-c)+\delta\frac{\hat{v}+1}{2}\frac{(1-c+\hat{v})^2}{4(1+\hat{v})}$ is maximized. This leads to $x = \frac{\delta(1-\delta)-(2-\delta-\frac{\delta^2}{2})c}{4-3\delta}$ which is positive if R2 is cheap and δ of some intermediate value. Otherwise, P must choose $x^1 = (0,c)$ and thus $\hat{v} = \frac{\delta}{2-\delta}(1-c)$, $x_1^2 = \frac{1-\delta}{2-\delta}(1-c)$, and $EU_P\left((0,c),(\frac{1-\delta}{2-\delta}(1-c),c)\right) = \frac{2(1-\delta)+\delta c}{2(2-\delta)}(1-c) + \frac{\delta}{2}(\frac{1-c}{2-\delta})^2$. P's second option would be to propose $x^1 = FAIL$ and $x^2 = (0,c)$. This would lead to $EU_P(FAIL,(0,c)) = \frac{\delta}{2}(1-c)$ which can be shown to be smaller than $EU_P\left((0,c),(\frac{1-\delta}{2-\delta}(1-c),c)\right)$ and thus smaller than $EU_P\left((\frac{\delta(1-\delta)-(2-\delta-\frac{\delta^2}{2})c}{4-3\delta}, c),(x_1^2,c)\right)$. □

This completes the proof of Proposition 2. □

Appendix A.2. Proof of Proposition 3

In sub-period $2a$, R1 will be offered $x = 0$ and he votes Y. That is because after a N-vote, P approaches R2 and secures an agreement with certainty (by offering c). In sub-period $1b$, after R1 has voted N, P will offer c to R2 if $(1-p)(1-c)+p\delta \geq \delta$, i.e., if $c \leq 1-\delta$.

Suppose $c \leq 1-\delta$. Then, in sub-period $1a$, R1 votes Y on an offer x if $x + v \geq (1-p)v + p\delta v$, i.e., if $x \geq -(1-\delta)pv = z$ or $v \geq \hat{v} = -\frac{x}{(1-\delta)p}$. So R1 votes Y with probability $\frac{1+\frac{x}{(1-\delta)p}}{2}$, given that $\frac{x}{(1-\delta)p} < 1$ (otherwise R1 votes Y for sure). P, in this case, chooses x such that $\frac{1+\frac{x}{(1-\delta)p}}{2}(1-x)+\frac{1-\frac{x}{(1-\delta)p}}{2}((1-p)(1-c)+p\delta)$ is maximized, leading to $x = \begin{cases} \frac{(1-p)c}{2} & p > \frac{c}{2-2\delta+c} \\ (1-\delta)p & otherwise \end{cases} = \min\{\frac{(1-p)c}{2}, (1-\delta)p\}$.

Suppose $c > 1-\delta$. Then, in sub-period 1a, R1 votes Y on an offer x if $x+v \geq \delta v$, i.e., if $x \geq -(1-\delta)v = z$ or $v \geq \hat{v} = -\frac{x}{(1-\delta)}$. R1 votes Y with probability $\frac{1+\frac{x}{(1-\delta)}}{2}$ and P thus maximizes $\frac{1+\frac{x}{(1-\delta)}}{2}(1-x) + \frac{1-\frac{x}{(1-\delta)}}{2}\delta$, leading to $x = 0$.

This completes the proof of Proposition 3. □

Appendix A.3. Proof of Proposition 4

First, recall that the sequence in which responders are approached does not matter. In Period 2, responders' announcements are truthful, exactly as in the "reliable" case. Therefore, Lemma 3 still holds when R1 has voted N in Period 1. Additionally, we must consider the scenario where R1 has voted Y, yet a Period 1 agreement has not been reached.

Lemma 6. *After R1 has voted Y in Period* 1*, P, under unanimity rule, proposes* $x^2 = (0, c)$*.*

Proof. After a R1-Y-vote in Period 1, P believes R1's valuation of agreement to be distributed on $v \in [\hat{v}, 1]$. So, given $x^2 = (x, c)$, R1's probability of voting Y is $\frac{1+x}{1-\hat{v}}$ if $x < -\hat{v}$ and 1 otherwise (he votes Y whenever $v > -x$). Given that $x < -\hat{v}$, P's Period 2 utility amounts to $EU_P(x^2 = (x, c)) = \frac{1+x}{1-\hat{v}}(1 - x - c)$. Maximizing this over x leads to $x = -c/2$. Hence, the optimal (and feasible) offer is $x^2 = (0, c)$ in all cases. □

Lemma 7. *Given* $x^1 = (x, c)$*, R1 votes Y iff* $v \geq \hat{v} = \begin{cases} \frac{\delta(1-c)-2(1-p)x}{2(1-p)(1-\delta)+\delta} & \hat{v} \in [0, 1-c] \\ \frac{\delta(1-c)-2(1-p)x}{2(1-p)-\delta} & \hat{v} \in [c-1, 0] \end{cases}$.
R1's Period 1 *price thus amounts to* $z(v) = \begin{cases} -(1-\delta)v + \frac{\delta(1-c-v)}{2(1-p)} & \hat{v} \in [0, 1-c] \\ -v + \frac{\delta(1-c+v)}{2(1-p)} & \hat{v} \in [c-1, 0] \end{cases}$,
which is increasing in p*.*

Proof. Voting Y gives R1 an expected utility of $(1-p)(x+v) + p\delta \max\{v, 0\}$. Voting N leads to an expected utility of $\delta \begin{cases} \max\{v, 0\} & \hat{v} > 1-c \\ \max\left\{v + \frac{1-c-\hat{v}}{2}, 0\right\} & \hat{v} \in [c-1, 1-c] \\ 0 & \hat{v} < c-1 \end{cases}$. So R1 is indifferent between Y and N if $v = \hat{v}$ and $(1-p)(x+v) + p\delta \max\{v, 0\} = \delta \begin{cases} \max\{v, 0\} & \hat{v} > 1-c \\ \max\left\{v + \frac{1-c-\hat{v}}{2}, 0\right\} & \hat{v} \in [c-1, 1-c] \\ 0 & \hat{v} < c-1 \end{cases}$. From this we obtain the cutoff type $\hat{v} =$

$\begin{cases} -\frac{x}{1-\delta} & \hat{v} > 1-c \\ \frac{\delta(1-c)-2(1-p)x}{2(1-p)-(1-2p)\delta} & \hat{v} \in [0,1-c] \\ \frac{\delta(1-c)-2(1-p)x}{2(1-p)-\delta} & \hat{v} \in [c-1,0] \\ -x & \hat{v} < c-1 \end{cases}$. Note that we can neither be in the upper case (because $\hat{v} \leq 0$ contradicts $\hat{v} > 1-c$) nor in the lowest one (because $x \in [0, 1-c]$). Consequently, the cutoff type is $\hat{v} = \begin{cases} \frac{\delta(1-c)-2(1-p)x}{2(1-p)(1-\delta)+\delta} & \hat{v} \in [0,1-c] \\ \frac{\delta(1-c)-2(1-p)x}{2(1-p)-\delta} & \hat{v} \in [c-1,0] \end{cases}$. So R1's price is $z(v) = \begin{cases} -(1-\delta)v + \frac{\delta(1-c-v)}{2(1-p)} & \hat{v} \in [0,1-c] \\ -v + \frac{\delta(1-c+v)}{2(1-p)} & \hat{v} \in [c-1,0] \end{cases}$. Hence, the price of R1 is increasing in p. □

Lemma 8. *P's optimal Period* 1 *offer is*

$$x^1 = \left(\max\left\{ \frac{(1-\delta)\delta - (1-\delta)c(2+\delta-2p+2p\delta) - \frac{\delta^2}{2(1-p)}c}{4(1-p)(1-\delta)+\delta}, 0 \right\}, c \right).$$

Proof. P chooses x such that the following expression is maximized:

$$\frac{1-\hat{v}}{2}\left((1-p)(1-x-c) + p\delta \max\left\{ \frac{1}{1-\hat{v}}, 1 \right\}(1-c) \right) + \frac{\hat{v}+1}{2}\delta \frac{\hat{v} - \frac{1-c-\hat{v}}{2}}{\hat{v}+1}(1-c-\frac{1-c-\hat{v}}{2}).$$

Given that $\hat{v} > 0$, this leads to the following Period 1 offer:

$$x = \frac{(1-\delta)\delta - (1-\delta)c\,(2+\delta-2p+2p\delta) - \frac{\delta^2}{2(1-p)}c}{4(1-p)(1-\delta)+\delta} \text{ or } x = 0.$$

Given $\hat{v} < 0$, the only equilibrium solution is:

$$x = \frac{\delta(1-c)}{2(1-p)},$$

which is dominated by:

$$x = \max\left\{ \frac{(1-\delta)\delta - (1-\delta)c(2+\delta-2p+2p\delta) - \frac{\delta^2}{2(1-p)}c}{4(1-p)(1-\delta)+\delta}, 0 \right\},$$

leading to $\hat{v} > 0$. □

Appendix B. Additional results

In this section, we have a look at some additional insights directly resulting from our model.

Appendix B.1. The responder's price

Note that under unanimity rule, R1's behavior for intermediate values of v is of our main interest. If R1 likes agreement very much, he will vote Y anyway: risking a breakdown is not worth it. If R1 dislikes agreement very much, he will vote N anyway: The pie is not large enough to compensate him sufficiently and he can be sure that there will be no agreement without his consent. Let us thus concentrate on $v \in (-1, 1-c)$ in the following. The following result is equivalent to Result 1.

Result. *R1 is (weakly) more willing to vote Y under majority rule than under unanimity rule because:*

- *If R1 negatively values reaching an agreement ($v < 0$): R1's Period 1 price under majority rule* $z(v)^m = \begin{cases} -(1-\delta)pv & c \leq 1-\delta \\ -(1-\delta)v & c > 1-\delta \end{cases}$ *is lower than his price under unanimity rule* $z(v)^u = \frac{\delta(1-c)-(2(1-p)(1-\delta)+\delta)v}{2(1-p)}$.
- *If R1 positively values reaching an agreement ($v \geq 0$): R1, in Period 1, votes Y on any offer under majority rule.*

Proof. Given that $v < 0$, R1's price under unanimity rule is larger than his price under majority rule: If $c > 1-\delta$, then $\frac{\delta(1-c)-(2(1-p)(1-\delta)+\delta)v}{2(1-p)} > -(1-\delta)v \Leftrightarrow \delta(1-c-v) > 0$. If $c \leq 1-\delta$, then $\frac{\delta(1-c)-(2(1-p)(1-\delta)+\delta)v}{2(1-p)} > -(1-\delta)pv \Leftrightarrow \delta(1-c-v) > 2(1-p)^2(1-\delta)v$. Both cases are true because $1-c-v > 0$ as $v < 0$. □

Appendix B.2. The probability of an overall breakdown

is the same as the probability of not finding an agreement and depends crucially on the voting rule that is applied. The following result is equivalent to Result 2.

Result. *Under majority rule, the probability of a breakdown is lower than under unanimity rule.*

Proof. Under majority rule: If $c > 1-\delta$, R2 is offered zero in Period 1 and R1 votes N with probability $\frac{1}{2}$ (if $v \geq 0$). Hence, breakdown occurs with probability $\frac{1-\delta}{2}$. If $c \leq 1-\delta$, R2 is offered c in Period 1 and votes N with probability p. R1, in subperiod $1a$, is offered $\tilde{x}$ and votes Y if $v \geq -\frac{\tilde{x}}{(1-\delta)p}$. So the probability of breakdown is even smaller than in the case where $c > 1-\delta$, i.e., it is $\in [0, \frac{1-\delta}{4}]$. More precisely, it is zero if $p < \frac{\tilde{x}}{1-\delta}$ and would be $\frac{1-\delta}{4}$ if $\tilde{x} = 0$ and $p = \frac{1}{2}$.

Under unanimity rule: there is breakdown with probability $\frac{2-\delta c}{2(2-\delta)}((1-\delta) + \delta\frac{1-\delta c+c}{2-\delta c}) = \frac{1}{2}$ if R2 is reliable ($p = 0$). If $p > 0$, then - as shown before - not only R2 is less likely to vote Y but also R1's willingness to vote Y shrinks. Consequently, the probability of breakdown must be weakly larger than in the "reliable" case. Consequently, the probability of breakdown is $\geq \frac{1}{2}$ under unanimity rule and $< \frac{1}{2}$ under majority rule. □

Appendix B.3. Welfare implications

Result. *The application of unanimity rule leads to more inefficient breakdowns and less inefficient agreements than the application of majority rule.*

Proof. Suppose $v < c-1$. Then having a breakdown is welfare optimizing. Whereas under unanimity rule there sure is a breakdown, it is still quite likely that there is an agreement under majority rule. Hence unanimity rule should be preferred over majority rule from a welfare perspective.

Suppose $v > c-1$. Then it follows directly from Result 2 that an inefficient breakdown is more likely under unanimity rule than under majority rule. □

Note that under unanimity rule, R1 might vote N on positive Period 1 offers even if he positively values agreement. That is, a breakdown might even occur in cases where ending up with a breakdown is very inefficient.

Example 1. *Take $c = \varepsilon$, $\delta = 1/2$, $v = 1/3 - \varepsilon$, and $p = 0$, with ε close enough to zero. In this scenario, an agreement results in a net benefit of $4/3 - 2\varepsilon$ compared to a breakdown. Still, under unanimity rule, a breakdown occurs half of the times (as R1 votes N on $x^1 = (0, \varepsilon)$). The (possible) Period 2 offer is $x^2 = (\frac{1-\varepsilon}{3}, \varepsilon)$. So expected profits are $\frac{1}{3} - \frac{2}{3}\varepsilon$ for R1 and $\varepsilon/2$ for R2.*

In contrast, the application of majority rule leads to a certain Period 1 agreement ($x^1 = (0, \varepsilon)$) and thus to expected profits of $\frac{1}{3} - \varepsilon$ for R1 and ε for R2. Notably, P is strictly better off in the latter case.

Thus, even in the most extreme scenario where the application of unanimity rule results in many inefficient and costly breakdowns, opting for majority rule does not lead to a Pareto improvement. This might change if there is asymmetry in the expected continuation probability δ.

Appendix C. Extensions

Now, let us turn to some extensions of the presented model.

Appendix C.1. Larger coalitions in a finite player game

Extension 1. Suppose both responders are of the unreliable cutoff type, with $c < \frac{1}{2}$ and $p < \frac{1}{2}$. Then, P will decide between three options: First, not offering anything in Period 1 and just "waiting" for Period 2 to be reached, leading to $EU = \delta(1-c)$. Second, approaching one of the responders and offering him c, leading to $EU = (1-p)(1-c) + p\delta(1-c)$. Third, offering c to both responders, resulting in $EU = (1-p^2)(1-2c) + p^2\delta(1-c)$.

P prefers the latter option over the first one if $\delta < \frac{1-2c}{1-c}$ and the latter over the second one if $c < \frac{p(1-p)(1-\delta(1-c))}{1+p-2p^2} = \frac{p(1-\delta(1-c))}{1+2p}$. Hence, P tries to form an oversized coalition (OSC) if

$$c < \frac{p\,(1-\delta(1-c))}{1+2p} \text{ and } \delta < \frac{1-2c}{1-c}.$$

Extension 2. Suppose that there are two additional responders (i.e., five players in total) and a simple majority rule is applied. Thus, there is an agreement if at least two of the four responders vote Y.[39] Suppose that two of the four responders (R2 and R3) are of the unreliable robot type, just like R2 in the original model: They both announce a Y-vote whenever offered at least c but then still end up voting N with probability $p < \frac{1}{2}$. R4 is of a reliable robot type but too expensive to be convinced of a Y-vote (e.g., he votes Y only if offered more than one). R1 is the responder of interest, like in the original model. However, R1's valuation of agreement now is continuously distributed on $[-1, 0]$.

In Period 2, all responders are reliable. P will approach R1 first, make him a zero offer, and R1 votes Y. That is because R1 knows that a N-vote would lead to a certain agreement with R2 and R3, while himself being excluded.

In Period 1, R1 will again be approached first:

- Suppose R1 has voted N. Then, P will have to decide whether she wants to offer c to R2 and R3 or just wait for the next period to be reached. The former is more beneficial than the latter if $(1-p)^2(1-2c)+(1-(1-p)^2)\delta(1-c) > \delta(1-c)$, i.e., if $c < \frac{1-\delta}{2-\delta}$ (for $\delta = 0$, this holds true because $c < 1/2$ is assumed).

[39] Recall that P is "assumed" to always vote Y.

- Suppose R1 has voted Y on the offer x. Then, P must decide whether she wants to offer c to both R2 and R3 or just to R2.[40] The former is more beneficial than the latter if $(1-p^2)(1-x-2c)+p^2\delta(1-c) \geq (1-p)(1-x-c)+p\delta(1-c)$, i.e., if $x \leq (1-\delta)(1-c) - \frac{1+p}{p}c$ (for $\delta = 0$, the condition is $x \leq 1 - \frac{1+2p}{p}c$).

Thus, building an OSC can only be optimal if responders are sufficiently unreliable (p not too small), responders not too costly (c small), and the continuation probability not too large (δ small): With reliable responders a MWC is optimal. If responders are expensive, an OSC becomes too expensive for P. If the next round is reached with high probability, then P is better off by risking no agreement today and building a MWC tomorrow.

For simplicity, suppose in the following that Period 2 will never be reached, i.e., $\delta = 0$. We will now look at the cases where the Period 1 offer x is small, such that P proposes an OSC after a Y-vote of R1, i.e., where $x \leq 1 - \frac{1+2p}{p}c$. Note that an offer larger than this is not even feasible for P if c is sufficiently small. Thus, it remains to be shown that P indeed wants to make a positive offer to R1 in such cases where $\delta = 0$ and c is not too large.[41]

Suppose R1 faces an offer $x \leq 1 - \frac{1+2p}{p}c$ in Period 1a. Then, he votes Y whenever $(1-p^2)(x+v)+p^2 \geq (1-p)^2 v + (1-(1-p)^2)$, i.e., if

$$v \geq -\frac{(1+p)x}{2p} \equiv \hat{v}.$$

R1 thus votes Y with probability $-\hat{v}$ and P therefore chooses her optimal offer x in order to maximize $-\hat{v}(1-p^2)(1-x-2c)+(1+\hat{v})(1-p)^2(1-2c)$. This leads to an optimal offer of

$$x = \frac{p}{1+p}(1-2c),$$

which is indeed smaller than $1 - \frac{1+2p}{p}c$ if $c < \frac{p}{1+3p}$ (otherwise we would have a corner solution with $x = \max\{0, 1 - \frac{1+2p}{p}c\}$).

[40] Waiting for the next round to be reached is on equilibrium path not optimal as she would then not have offered x to R1 in the first place.

[41] If offering $x > 1 - \frac{1+2p}{p}c$ would be feasible and would be offered, R1 votes Y whenever $(1-p)(x+v) \geq (1-p)^2 v$, i.e., if $v \geq -\frac{x}{p} \equiv \hat{v}$. P thus maximizes $-\hat{v}(1-p)(1-x-c)+(1+\hat{v})(1-p)^2(1-2c)$, leading to $x = \frac{c+(1-2c)p}{2}$. This is only larger than $1 - \frac{1+2p}{p}c$ if c is large. Hence, as discussed above, when focusing on small c, the case $x > 1 - \frac{1+2p}{p}c$ becomes irrelevant for on-equilibrium path behavior.

Note that R1's price is

$$z = -\frac{2p}{1+p}v$$

and thus increasing in p. In line with the original model, the unreliability of the other responders makes R1 more expensive. The optimal proposal to R1 is also increasing in p, i.e., R1 is being compensated for the increased risk of proposal failure.

Note that if p is very small, then offering $x \leq 1 - \frac{1+2p}{p}c$ does not work and P would be better off by proposing a MWC. Thus, the unreliability of the other responders makes it indeed likely that P forms an OSC. Particularly, this is the case if δ and c are small, i.e., if it is unattractive for P to wait for another round of bargaining and if the unreliable responders are not too expensive.

Appendix C.2. The continuation probability δ

The continuation probability δ is one of the main drivers for voting N under unanimity rule. It increases the *positive signaling value from voting N* and thus makes R1 more expensive. Consequently, incorrect beliefs about δ can lead to (too) many rejections of Period 1 proposals. For simplicity, assume $p = 0$ in the following.

Lemma 9. *Given that R1 overestimates the probability to go into a second round of bargaining (i.e., $\hat{\delta}_R > \delta$[42]), applying majority rule instead of unanimity rule might be a Pareto improvement.*

Proof. Take $c = \varepsilon$, $\hat{\delta}_R = 3/4 > \delta = 1/2$ and $v = 3/5 - \varepsilon$, with ε close enough to zero. Under unanimity rule, P proposes $x^1 = (0, \varepsilon)$, R1 votes N (as $v < \hat{v} = \frac{\hat{\delta}_R}{2-\hat{\delta}_R}(1-\varepsilon)$), and if Period 2 is reached (with probability $\frac{1}{2}$), P proposes $x^2 = (\frac{1-\hat{\delta}_R}{2-\hat{\delta}_R}(1-\varepsilon), \varepsilon)$. So expected profits are $\frac{1}{2}(\frac{1}{5}(1-\varepsilon) + \frac{3}{5} - \varepsilon) = \frac{2}{5} - \frac{3}{5}\varepsilon$ for R1 and $\varepsilon/2$ for R2 under unanimity rule.

Under majority rule, there is an immediate agreement on $x^1 = (0, \varepsilon)$ and profits are $\frac{3}{5} - \varepsilon$ for R1 and ε for R2. So here switching from unanimity rule to majority rule is Pareto efficient (P is always better off under majority rule). □

[42] $\hat{\delta}_R$ is defined as the (incorrect) belief of R1 about the continuation probability δ.