Proto-coalition bargaining and the core

Breitmoser, Yves

EUV Frankfurt (Oder)

14 September 2010

Proto-coalition bargaining and the core

Yves Breitmoser*

EUV Frankfurt (Oder)

September 14, 2010

Abstract

In the proto-coalition model of government formation, formateur F appoints a proto-coalition and asks its members whether to start negotiating a coalition contract. If all accept, then the proto-coalition forms and starts negotiating; otherwise a caretaker government assumes office. I extend this model by allowing F to revise the chosen proto-coalition after rejections, that he states pre-conditions for the subsequent negotiations, and that F's opponents may publicly pre-commit to accept/reject certain proposals. The set of equilibrium outcomes is identified as the core if F's opponents can pre-commit and as the convex hull of the core if they cannot pre-commit credibly. This extended model eliminates two flaws of the standard model: it explains why F cannot always install his favored coalition (whatever the status quo) and why "important" coalition members may have more bargaining power in the subsequent negotiations than others.

JEL–Codes: C72, C78, D72

Keywords: coalition formation, non-cooperative bargaining, core

*I thank Friedel Bolle, Frank Bönker, and Jonathan Tan for helpful comments. Financial support of the DFG, project no. BO $747/10-1$, is gratefully acknowledged. Address: Postfach 1786, 15207 Frankfurt(Oder), Germany, email: breitmoser@euv-frankfurt-o.de, Telephone/Fax: +3355534 2291/2390.

1 Introduction

The process of government formation is central in multiparty democracies. It determines the composition of the government coalition, which in turn sets the political agenda and essentially passes the legislation. As a result, the institutional details of elections and government formation affect income tax rates (Austen-Smith, 2000), economic policy (Persson, 2002), fiscal policy (Persson and Tabellini, 2004), and many related aspects. To analyze the institutional details accurately, increasingly refined models of government formation have been developed in the literature (see for example Diermeier, 2006, for a survey). The model that seems to capture the empirical regularities best is the model of "proto-coalition bargaining" introduced by Diermeier and Merlo (2000) and Baron and Diermeier (2001): The government formateur F chooses a proto-coalition c, and if all members of c accept, then a coalition contract is negotiated (using the model of Merlo and Wilson, 1995, 1998).[1] Otherwise, a default (caretaker) government assumes office.

This model is sufficiently flexible to account for the occurrence of minority, minimal-majority, and surplus governments (by varying the default payoffs), and if extended to account for out-of equilibrium phenomena, its parameters can be estimated based on real-world data. The estimated structural models, in turn, have been used in counterfactual policy experiments to study institutional and constitutional design (Diermeier et al., 2002, 2003, 2007). In this paper, I analyze an extension to the model of proto-coalition bargaining that will be shown to account for two additional phenomena that the standard model fails to explain in equilibrium: (i) F cannot always install his preferred coalition (even if the default payoffs are low) and (ii) coalition members that are "important" (to be specified) have more bargaining power in the subsequent negotiations than other parties. Thus, the range of outcomes compatible with the proto-coalition model is widened substantially, and hence the extended model allows more precise analyses of institutional and constitutional design.

[1] The main alternative models of coalitional bargaining, random proposer bargaining (Baron and Ferejohn, 1989; Eraslan, 2002; Kalandrakis, 2006) and demand commitment bargaining (Winter, 1994; Cardona-Coll and Mancera, 2000) apply less well to government formation (as opposed to parliamentary bargaining, for example), as the identity of the formateur is not constant in these models.

The main extension that I propose is as follows. If one of the parties in the proto-coalition chosen by F refuses to start negotiating a coalition contract, then the standard model posits that a caretaker government is instated. In most cases, the government formateur F may actually make a second proposal, or a third one if necessary, and more generally modern means of communication render the existence of remote time lines strategically irrelevant. Hence, government formation is more aptly modeled as a game with infinite time horizon, at least as long as the time line is considered "remote," as opposed to a finitely repeated game or even a one-round game. Thus, action and reaction between formateur and potential coalition partners are strategically relevant already prior to the formation of the proto-coalition. An implication of this interaction is that F's opponents may reject the first proposal of F to force him into forming the proto-coalition of their choice, rather than sticking with his choice. The strategic relevance of the coalition preferences of F in relation to those of the other players is empirically obvious (as different parties have different political platforms and hence prefer to coalesce with different partners), but their relation is irrelevant in the one-round model. In contrast, the standard one-round model explains deviations from F's favorite choice by assuming either high default payoffs or bounded rationality of F (as in logit choice functions, see Diermeier et al., 2003). Allowing F to revise rejected proposals eliminates this neglect of the standard model.

Two other extensions to the standard model that I analyze are as follows. On the one hand, I allow F to offer ideological or distributive concessions during coalition formation (formalized as "pre-conditions" below), using which he may try to assemble his favored coalition. On the other hand, I allow F's opponents to pre-commit publicly to accept or reject certain proposals of F. The results can be summarized as follows. The extended proto-coalition game implements a "quasi-core" if F's opponents may pre-commit publicly and it implements a superset of the quasi-core if they cannot. These solution sets equate with the core and the convex hull of the core (respectively) under the standard assumption that the status quo is bad. This relationship between proto-coalition bargaining and the core is novel and will be established under fairly general conditions. I then analyze several classes of example games to illustrate the main characteristics of the extended model, and amongst others these examples illustrate why F cannot always form his favored coalition (essentially, when it is not the favored coalition of the other players) and why F may have to give up bargaining power during coalition formation.

The notation is introduced in Section 2. Section 3 analyzes the game where F's opponents cannot pre-commit credibly, Section 4 analyzes the game where they can. Section 5 analyzes the example games and Section 6 concludes.

2 The base model

Definitions The formateur is denoted as F. The remaining players are denoted as $i \in N = \{1, \ldots, n\}$. The game proceeds in rounds. In each round, F chooses a proto-coalition $c \subseteq N_F := N \cup \{F\}$ such that $F \in c$ and pre-conditions $r \in \mathcal{R}$, where $|\mathcal{R}| < \infty$. If at least one $i \in c$ does not agree to negotiate under pre-conditions r, then a new round begins. Otherwise, the proto-coalition forms and multilateral negotiations start. The protocol of these negotiations is left open in the general analysis.[2] Regardless of the protocol, the expected payoffs from the negotiations are common knowledge in equilibrium. They are denoted as $v_i(r, c)$ for all $i \in N_F$ and assumed to be non-degenerate, see Eq. (1) below. Given the discreteness of the pre-conditions, non-degeneracy is implied in most models of intra-coalitional bargaining. Examples are discussed in Section 5. In turn, discreteness of $\mathcal{R}$ is required to express categorical pre-conditions, e.g. whether new nuclear power plants may be constructed, whether taxes may be raised, or whether one participates in international military campaigns. A continuity assumption would be inappropriate in such cases, as pre-conditions prior to government formation do generally not concern quantitative details such as size of power plants or actual tax rates.

$C(r) \subseteq \mathcal{P}(N_F)$ denotes the set of proto-coalitions to which pre-conditions r may be proposed; $\mathcal{P}(N_F)$ is the power set of N_F. In most applications, $C(r)$ would satisfy either *uniformity* (all pre-conditions are applicable to all decisive coalitions) or

[2] In this way, most reasonable models of the subsequent negotiations are treated in a unified manner, e.g. random or deterministic proposer models, models of distributive and ideological decisions, and different degrees of randomness of cake sizes. In such cases, the pre-conditions may concern relative weights for proposal making, rights of first proposal, or restrictions of the bargaining outcome.

specificity (every $r \in \mathcal{R}$ is applicable to exactly one coalition). The following analysis treats these (and all other) cases in a unified manner. Without loss of generality, assume that if $C(r)$ is not a singleton, then the expected payoff v_i is independent of which proto-coalition $c \in C(r)$ accepts r. This comes without loss, since payoff relevance can be expressed by refining $\mathcal{R}$ as necessary. Thus, $v_i(r) \equiv v_i(r,c)$ for all $i \in N_F$ and all $c \in C(r)$, and non-degeneracy can be defined as

$$\forall r, r' \in \mathcal{R}\ \forall i \in N_F : v_i(r) \neq v_i(r'). \tag{1}$$

Future payoffs are discounted by $\delta \in (0,1)$, and the valuations of the status quo (which applies until agreement is reached) are denoted as $\tilde{v}_i$ for all $i \in N_F$. Note that the status quo need not be "bad" in the sense of Banks and Duggan (2006). To summarize, the game is defined as a tuple $\Gamma = \langle N, \mathcal{R}, C, (v_i), (\tilde{v}_i) \rangle$.

In addition, I use $r_1 \succsim_i r_2$ as a short-cut for $v_i(r_1) \geq v_i(r_2)$, including all obvious variations (e.g. $r_1 \succ_i r_2$), and the partial ordering "$\succ_c$" is used as follows.

$$\forall r_1, r_2 \in \mathcal{R} : \qquad r_1 \succ_c r_2 \quad \Leftrightarrow \quad \forall i \in c : \ r_1 \succ_i r_2. \tag{2}$$

Strategies The set of t-round histories is $H_t = \left(P_F \times \{0,1\}^N\right)^t$, with $t \geq 0$ and

$$P_F := \{(r,c) \mid r \in \mathcal{R} \text{ and } c \in C(r)\} \tag{3}$$

as the set of proposals that F can make. The set of all histories is $H = \cup_{t \geq 0} H_t$. F's strategy is a function $\tau_F : H \to P_F$, and for all $i \in N$, strategies are denoted as $\tau_i : H \times P_i \to \{0,1\}$, with

$$P_i := \{(r,c) \mid r \in \mathcal{R} \text{ and } i \in c \in C(r)\} \tag{4}$$

as the set of proposals addressing a proto-coalition including i. For example, $\tau_i(h)(r,c)$ is the probability that i accepts entering multilateral negotiations after history h under the proposal (r,c). I will characterize the set of (pure) equilibria that are semi-stationary in the following sense.

Definition 2.1 (Semi-stationarity)**.** Let $\mathbf{p}(h)$ denote the sequence of proposals $(r^t, c^t)_{t \geq 0}$ implied by the history of play $h \in H$ and define $\mathbf{P} = \cup_{h \in H} \mathbf{p}(h)$. Semi-stationary perfect equilibria τ are perfect equilibria (Selten, 1975) that are measurable with respect to set of proposal sequences $\mathbf{P}$, i.e. for all histories $h, h' \in H$,

$$\mathbf{p}(h) = \mathbf{p}(h') \quad \Rightarrow \quad \tau(h) = \tau(h'). \tag{5}$$

That is, the strategies may depend on the proposal sequence F made before (e.g. on which "concessions" F made), but they are independent of which player happened to reject a proposal made in the past. This notion of "semi-stationarity" is not to be confused with the (context dependent) alternative definitions of semi-stationarity provided by Rubinstein and Wolinsky (1985), Wooders (1998), and Kultti (2000). I relax the standard assumption of stationarity, as strict stationarity would imply that players ignore which proposals/concessions had been made before, which seems unrealistic in the context of government coalition. Fully unrestricted non-stationarity, in turn, does not allow me to express the intuitive notion of sincere voting (which is implied in semi-stationary perfect equilibria). Without semi-stationarity, one's payoff maximizing voting decision would depend on who (if any) is anticipated to reject the standing proposal, and this would not be "sincere."

3 Analysis of the base model

In a first step of the analysis, I characterize the outcomes that may be sustained in stationary equilibria. Under stationarity, F proposes the same pre-conditions r to the same proto-coalition c in every round. In equilibrium, (r,c) is accepted without delay and no player is better off deviating unilaterally. On the one hand this implies that preconditions are accepted only if no player benefits from delaying agreement. Hence, r must dominate the status quo for all $i \in c$, i.e. $r \in \mathcal{R}^*$ as defined next.

$$\mathcal{R}^* = \{r \in \mathcal{R} \mid \exists c \in C(r) : r \succsim_c \tilde{v}_i\} \tag{6}$$

On the other hand, if all players anticipate that some $r \in \mathcal{R}$ is going to be proposed in all future rounds round, and if there exists some (r',c') that all $i \in c'$ prefer to r, then F is better off deviating to propose (r',c') in any round. For, all $i \in c'$ will accept it when evaluating it in relation to their continuation payoffs $v_i(r)$. To summarize, the set of stationary equilibrium outcomes is the *quasi core* of Γ as defined next.[3]

$$QC(\Gamma) = \{r \in \mathcal{R}^* \mid \nexists (r', c') \in P_F \text{ such that } r' \succ_{c'} r\} \tag{7}$$

It is called "quasi core," because being "undominated in $\mathcal{R}$"[4] is not sufficient for the (quasi-) core property; in addition, the respective options have to dominate the status quo. It coincides with the core of the formateur bargaining game if the status quo is "bad," and in this case, $QC(\Gamma)$ is guaranteed to be non-empty.

Next, we characterize the outcome set under semi-stationarity. It contains all outcomes compatible with stationarity, and additionally allows for two effects. On the one hand, if a decisive coalition c' prefers negotiating under r' to some alternative (r,c), then c' accepts r' if F "threatens" to go for (r,c) otherwise. Such threats can be reiterated as long as c^k prefers r^k over (r^{k-1}, c^{k-1}) for all $k \geq 1$, as in

$$r^k \succ_{c^k} r^{k-1} \succ_{c^{k-1}} \cdots \succ_{c^2} r^1 \succ_{c^1} r^0, \tag{8}$$

Any r^k that is sustained by such a sequence can result in equilibrium if the seed (r^0, c^0) is an equilibrium outcome. If (r^0, c^0) is to be sustained in a stationary continuation equilibrium, however, then it is an element of the quasi core, and its very definition thus rules out that it could be the seed of any such sequence. It follows that the continuation equilibrium inducing (r^0, c^0) cannot be stationary. A second issue with the credibility of iterating threats as in (8) is that F must not be able to deviate profitably to a proposal other than the one listed next in the sequence.

To illustrate how these issues can be resolved, let me first illustrate how elements of $\mathcal{R}^*$ that F prefers to some element of the quasi core can result under semi-stationarity without iterating threats. Define

$$\overline{QC}(\Gamma) := \{r \in \mathcal{R}^* \mid \exists r' \in QC(\Gamma): \ r \succsim_F r'\}, \tag{9}$$

pick any $r \in \overline{QC}(\Gamma) \setminus QC(\Gamma)$, and choose $r' \in QC(\Gamma)$ such that $r \succsim_F r'$. These additional elements $r \in \overline{QC}(\Gamma) \setminus QC(\Gamma)$ are potential seeds of iterated threats as in (8). An equilibrium resulting in r along the path is: (i) F proposes r if he has never deviated

[3]This definition the quasi core should not be confused with alternative ones, as in Shapley and Shubik (1966) and Shimomura (1997).

[4]*r dominates r'* if $r \succ_c r'$ for some $c \in C(r)$ and r is *undominated in $\mathcal{R}$* if no $r' \in \mathcal{R}$ dominates r

from proposing r in the past, (ii) F proposes r' otherwise, and (iii) all players vote sincerely. Hence, if F ever deviates from proposing r, then the continuation outcome becomes r', and since r' is undominated in $\mathcal{R}$, no pre-conditions but r' would be accepted when r' is the continuation outcome. Thus, when F deviates from c, then r' results indeed, and since $r \succsim_F r'$, F is indeed best off proposing r initially. In this way, r results in equilibrium although r is not undominated in $\mathcal{R}$. The following shows how this generalizes to sequences of iterated threats as in (8).

Lemma 3.1. *Generically, there exists $\underline{\delta} \in (0,1)$ such that $r \in \mathcal{R}$ can result in a semi-stationary perfect equilibrium for all $\delta \in \big(\underline{\delta}, 1\big)$ if there exists $(r^k, c^k)_{k \leq K}$ for some $K \in \mathbb{N}_0$ such that $r = r^K$ and*

$$r^K \succ_{c^K_{-F}} r^{K-1} \succ_{c^{K-1}_{-F}} \cdots \succ_{c^2_{-F}} r^1 \succ_{c^1_{-F}} r^{k'} \quad \textit{for some } k' \leq K, \tag{10}$$

and there exists $r^* \in QC(\Gamma)$ such that $r^k \succsim_F r^*$ for all $k \leq K$.

Proof. First, I show that $r^* \in QC(\Gamma)$ can result in a (stationary) perfect equilibrium. Define the strategy profile τ where (i) F invariantly proposes r^* and (ii) all opponents vote sincerely. Formally, for all $h \in H$,

$$\tau_F(h) = (r^*, c^*), \tag{11}$$

$$\forall i \in N \quad \forall (r,c) \in P_i : \quad \tau_i\big(h, (r,c)\big) = \begin{cases} 1, & \text{if } r \succsim_i r^*, \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

By construction, it is a stationary strategy profile that results in r^* along the path of play. The voting decisions $\tau_i\big(h, (r,c)\big)$ are compatible with perfection, as $v_i(r^*)$ is i's continuation payoff in case (r,c) is rejected. To see that F's proposals $\tau_F(h)$ are compatible with perfection, too, note first that $r^* \in QC(\Gamma)$ implies $\nexists r' \in \mathcal{R}$ such that a decisive coalition prefers r' to r^*. Hence, and given the voting decisions of the $i \in N$, F is strictly best off proposing r^*. Any $r' \succ_F r^*$ that F prefers will be accepted only with infinitesimal probability, and hence proposing any $r' \succ_F r^*$ is not optimal if $\delta < 1$. Proposing options r' with $r^* \succ_F r'$ cannot be optimal regardless of their acceptance probabilities, and for appropriately constructed tremble probabilities, proposing (r^*, c^*) is strictly more profitable than proposing any (r', c') where $r' \sim_F r^*$. Combined, this implies that τ is a perfect equilibrium.

Second, consider r as claimed, define $(r^k,c^k)_{k\leq K}$ for $K \in \mathbb{N}_0$ and $r^* \in QC(\Gamma)$ such that $r^k \succsim_F r^*$ for all $k \leq K$. Consider the following strategy profile.

1. F proposes the options in the order $\left(r^K,c^K\right),\ldots,\left(r^1,c^1\right),\left(r^{k'},c^{k'}\right),\left(r^{k'-1},c^{k'-1}\right),\ldots$ until one is accepted. If F ever deviates from this sequence, then the continuation strategies are the stationary equilibrium sustaining r^* (see above).

2. F's opponents vote sincerely (anticipating proposals as defined in point 1).

By construction, this strategy profile is semi-stationary, and the argument establishing perfection is similar to the one above. Sincere voting implies that any proposal (r^k,c^k) along this sequence would be accepted by the respective proto-coalition. This in turn implies that any proposal deviating from this sequence would be rejected by at least one player of any proto-coalition (since the next proposal would be r^* which is undominated in $\mathcal{R}$). Given this, F is best off sticking to the outlined proposal sequence until a proposal is accepted. □

Lemma 3.1 implies that a sequence need not be seeded in some $r^0 \in \overline{QC}$ for being sustainable in equilibrium. Alternatively, the sequence may be circular. The following result completes the characterization of equilibrium outcomes under semi-stationarity. It shows that it is not necessary that the "fall-back result" r^* is in $QC(\Gamma)$, i.e. it is not necessary that F's opponents prefer it to the status quo. The solution set will be defined using a notion of self-generation as it is known from dynamic programming.

Definition 3.2 (Self-generation)**.** $r \in R$ is *enforceable on* $R' \subseteq \mathcal{R}$ if $r \succ_{c_{-F}} r'$ for some $c \in C(r)$ and some $r' \in R'$. The set $R' \subseteq \mathcal{R}$ is *self-generating* in Γ if

1. for all $r \in R'$: r dominates the status quo $\tilde{v}$ or r is enforceable on R', and

2. $\min R'$ is undominated in $\mathcal{R}$ and $\min R' \succ_F \tilde{v}$

using $\min R' := \arg\min_{r' \in R'} v_F(r')$. The largest self-generating set is denoted as $SG(\Gamma)$.

As the most important special case, consider the standard assumption that the status quo is bad, i.e. $v_i(r) > \tilde{v}_i$ for all $i \in c \in C(r)$ and all r. Then, the largest self-generating set simply convexifies the core under $\succsim_F$, i.e. in this case, it simplifies to

$$\text{Bad status quo:}\quad SG(\Gamma) = \{r \in \mathcal{R} \mid \exists r', r'' \in QC(\Gamma) \text{ such that } r' \succsim_F r \succsim_F r''\}.$$

This will be illustrated and further discussed in Section 5 (e.g. Lemma 5.2). The next result establishes $SG(\Gamma)$ as the solution set in general.

Proposition 3.3. *Generically, there exists $\underline{\delta} \in (0,1)$ such that r can result in a semi-stationary perfect equilibrium for all $\delta \in \big(\underline{\delta}, 1\big)$ if and only if $r \in SG(\Gamma)$.*

Proof. The proof of Lemma 3.1 is adapted straightforwardly to show that $r^* \in SG(\Gamma)$ is sufficient. To establish this point, consider the following strategy profile.

1. F proposes (r^k, c^k) in order until one is accepted; if F ever deviates, then the continuation equilibrium sustaining $r_{\min} = \min SG(\Gamma)$ is adopted

2. F's opponents vote sincerely in all cases

As above, sincere voting implies that the players accept any proposal along the sequence (since every proposal dominates its respective successor), and they reject any other proposal (since $r_{\min}$ is the continuation outcome, which is undominated in $\mathcal{R}$). Anticipating this, F is best off sticking to the proposal sequence.

It remains to show that $r^* \in SG(\Gamma)$ is necessary. For contradiction, assume a semi-stationary perfect equilibrium τ exists that results in some $r^* \notin SG(\Gamma)$. Define $R' \subseteq \mathcal{R}$ as the set of options that are accepted by some proto-coalition after some history of play under τ. I show that R' must be self-generating in the sense of Def. 3.2, which yields the contradiction.

First, assume that $\min R'$ is not undominated in $\mathcal{R}$, and consider any subgame where $r' := \min R'$ is supposed to result. Assume that F deviates (in this subgame) from proposing r' toward proposing any r'' that dominates r'. In case r'' is accepted according to τ, then F benefits from this unilateral deviation. In case r'' is not accepted according to τ, then either r' is the continuation outcome or some $r''' \succ_F r'$ (since $r' = \min R'$ and the valuations are generic). The former case contradicts perfection in that the proto-coalition does not vote sincerely, and the latter implies that F benefits by deviating toward r'' in the considered subgame. *Second*, assume $\min R' \not\succsim_F \tilde{v}$. In this case, F is best off deviating from τ in subgames where F is supposed to propose $\min R'$ (anything worse than $\min R'$, in F's eyes, cannot result from this deviation, and $\min R'$ itself would be delayed, to the benefit of F). *Third*, assume that some $r' \in R'$ neither dominates the status quo $\tilde{v}$ nor is enforcable on R'. Consider any subgame where r' results according to τ, i.e. is both proposed by F and accepted by the respective proto-coalition, and let r'' denote the continuation outcome if it would be rejected by the proto-coalition (under semi-stationarity, r'' is independent of the individual voting decisions). If $r' = r''$, then r' is accepted in a perfect equilibrium only if it dominates the status quo. Since r' does not dominate the status quo by assumption, $r' \neq r''$ must apply. Such r' is accepted by a proto-coalition (in a perfect equilibrium, for $\delta \approx 1$) only if r' dominates r'', i.e. if it is enforcable on R' (the contradiction). $\square$

4 Players can publicly pre-commit

Next, we analyze the extension of the above game where F's opponents may announce either *negative* or *positive* pre-commitments prior to F's proposal. A negative pre-commitment is one where i commits to reject negotiating under the respective pre-conditions, and a positive pre-commitment is one where i commits to accept negotiating under these pre-conditions. The exact move structure is as follows.

Definition 4.1 (Extended move structure)**.** The game proceeds in rounds as above. In each round, first the non-formateur players announce pre-commitments with respect to any $r \in \mathcal{R}$, second F chooses (r,c), i.e. proposes coalition c to negotiate under pre-conditions r, and third all $i \in c$ vote on (r,c) to the degree they are uncommitted. The proto-coalition is formed if all $i \in c$ accept; otherwise a new round begins.

I assume that pre-commitments have to be renewed when a proposal of F had been rejected, i.e. after each round. Any assumption of finite duration is outcome equivalent to this one-round assumption. Alternatively, the game where pre-commitments can be made once and for all, prior to the first round, is outcome equivalent to a game with finite time horizon.

The additional notation can be kept brief. Primarily, $D \subseteq \{-1,0,1\}$ denotes the set of "directions" in which non-formateur players may pre-commit. By varying

D, four kinds of games can be distinguished. If $D = \{-1, 0, 1\}$, then both positive and negative pre-commitments are possible, if $D = \{0\}$, then pre-commitments are impossible, if $D = \{0, 1\}$, then only positive commitments are possible, and if $D = \{-1, 0\}$, then only negative commitments are possible. The game induced by $D = \{0\}$ has been analyzed above. The present section considers the remaining three games.

Strategy profiles are now pairs (κ, τ) with τ_i as defined above and $\kappa_i(h, r) \in D$ as the pre-commitment of $i \in N$ chosen after history h with respect to pre-condition $r \in \mathcal{R}$. For example, $\kappa_i(h, r) = 1$ implies that i pre-commits to accept r when it should be proposed by F. I consider perfect equilibria (κ, τ) in semi-stationary strategies.[5] The next result shows that in case positive pre-commitments are impossible, the possibility of negative pre-commitments is outcome irrelevant.

Proposition 4.2. *Assume positive pre-commitments are not possible, i.e.* $1 \notin D$. *Generically, there exists* $\underline{\delta} \in (0, 1)$ *such that for all* $\delta \in (\underline{\delta}, 1)$, *$r$ can result in equilibrium iff* $r \in SG(\Gamma)$.

Proof. First, I show that any $r^* \in SG(\Gamma)$ may result in equilibrium. Fix $r^* \in SG(\Gamma)$, $c^* \in C(r^*)$, and define a strategy profile sustaining (r^*, c^*) as well as $r_{\min} = \min SG(\Gamma)$ as in the proof of Prop. 3.3. That is, F's opponents do not pre-commit along the path of play, they vote sincerely in all cases, and F picks proposals from a sequence as above. Given Lemma 3.1, it remains to show that F's opponents cannot benefit by making (negative) pre-commitments. Consider an arbitrary subgame; let (r, c) denote its equilibrium outcome, and let (r', c') denote the respective continuation outcome. If any of the players $i \in c_{-F}$ pre-commits to reject (r, c), then F is best off following the proposal sequence nonetheless (since $r' \succsim_F r_{\min}$); hence negative pre-commitments are optimal under the constructed strategy profile only for players who would reject the respective proposal. Since the latter is ruled out by construction of the proposal sequence, no player can gain by deviating unilaterally to a negative pre-commitment.

Second, I show that only $r^* \in SG(\Gamma)$ may result in equilibrium. Assume the opposite and let $R' \subseteq \mathcal{R}$ denote the set of (continuation) outcomes of an equilibrium

[5] To clarify, players with pre-commitments to accept or reject specific proposals can tremble just as players with corresponding pure strategies can. Otherwise, perfection does not induce sincere voting of the uncommitted players.

τ contradicting this claim. The arguments showing that $\min R'$ need be undominated in $\mathcal{R}$ and that $\min R' \succ_F \tilde{v}$ follows equal those from Proposition 3.3. It remains to be contradicted that $\exists r \in R'$ that neither dominates the status quo $\tilde{v}$ nor is enforcable on R'. Let c denote the proto-coalition accepting r in the respective subgame, and let (r', c') denote the continuation result in case F proposes (r, c) but gets rejected. If $r \neq r'$, then $r \succ_{c_{-F}} r'$ would have to apply (since r would not be accepted otherwise), and since $r' \in R'$ by definition of R', r would then be enforcable on R'. Hence, $r = r'$ must hold true. Now, either (i) $r \not\succ_i \tilde{v}$ for some $i \in c_{-F}$, or (ii) $r \not\succ_F \tilde{v}$ must be satisfied. In case (i), the respective $i \in c_{-F}$ would gain by deviating unilaterally toward rejection of (r, c), which contradicts the assumption that r may result in equilibrium, and in case (ii), a contradiction to $\min R' \succ_F \tilde{v}$ results. $\square$

The result that the possibility of negative pre-commitments is outcome irrelevant may be surprising. The following explains the intuition. Announcing "knockout criteria" (i.e. negative pre-commitments) allows non-formateur players to prevent F from stating exorbitant pre-conditions. Now assume that a player pre-commits to reject the first element (r', c') from a sequence of (iteratively self-enforcing) proposals. As a result, F will have to settle with the next-best option, say (r, c). By assumption, all $i \in c'$ prefer (r', c') over (r, c). Hence, pre-committing negatively with respect to r' indeed would prevent it, but it does so only in circumstances where i would actually want F to ask for r'. Players do therefore not benefit from announcing knockout criteria in equilibrium, and the possibility do so is outcome irrelevant.

In contrast, consider the implications of positive pre-commitments. Their possibility induces competition reminiscent of auctions between the non-formateur players, and in a first induction step F seems to benefit from it. Assume for example that the equilibrium induces some (r, c) along the path of play, and that an option (r', c') exists that all members of the proto-coalition c' (including F) prefer to (r, c). If they pre-commit to accept r', then F will propose (r', c') instead of (r, c), and all of them are better off. Hence, no such (r, c) may result along the equilibrium path. However, this applies both in the beginning of the game, and more critically in any continuation equilibrium after any history h. As a result, the "sustaining" threats required to build complex self-generating sets are not credible anymore—i.e. the (r, c) thanks to which

c' accepts (r', c')—and for this reason, only singleton self-generating sets can now be sustained in equilibrium. This leads us back to the quasi core, and in turn, F cannot play the proto-coalitions off against one another if they can pre-commit positively.

Proposition 4.3. *Assume positive pre-commitments are possible, i.e.* $1 \in D$*. Generically, there exists* $\underline{\delta} \in (0,1)$ *such that for all* $\delta \in (\underline{\delta}, 1)$*, r can result in equilibrium iff* $r \in QC(\Gamma)$*.*

Proof. First again, I show that any $r^* \in QC(\Gamma)$ may result in equilibrium. Fix r^* and $c^* \in C(r^*)$, and let (κ, τ) denote a strategy profile as follows: (i) no pre-commitments are made, (ii) F makes the payoff maximizing proposal given the actual pre-commitments and anticipating (r^*, c^*) as the continuation outcome, and (iii) all $i \in N$ vote sincerely subject to standing pre-commitments. This strategy profile is stationary, hence also semi-stationary, and results in (r^*, c^*) along the path of play. Mutual optimality of both voting and proposal functions follows from their definitions. It has to be shown that the players are best off not to pre-commit after any history of play. F makes a proposal other than (r^*, c^*) iff an (r, c) exists such that $r \succ_F r^*$, and $k_{ir} = 1$ or $r \succ_i r^*$ for all $i \in c$. It follows that only positive pre-commitments may be payoff relevant for any $i \in N$, and they are payoff relevant only if they concern $r \in \mathcal{R} : r \succ_F r^*$ and $r \not\succ_i r^*$. Since the continuation payoff is (r^*, c^*) in all cases, such pre-commitments are at least weakly dominated in all subgames, and they are generally strictly dominated under full support. Hence, F's opponents are never best off deviating from the above strategy profile τ by making positive or negative pre-commitments.

Second, I show that only $r^* \in QC(\Gamma)$ may result in equilibrium. Assume the opposite, i.e. some $r \notin QC(\Gamma)$ may result in an equilibrium τ. *Case 1:* r is not undominated in $\mathcal{R}$. Hence, there exists (r', c') such that $r' \succ_{c'} r$. If all $i \in c'_{-F}$ would pre-commit to accept r', then F will be best off deviating to propose (r', c') in the corresponding subgame. For, since F cannot deviate profitably from proposing r along the equilibrium path (recall that τ is a semi-stationary equilibrium), F must be best off proposing $r' \succ_F r$ in the subgame where r' will be accepted almost surely due to pre-commitments. In turn, as each $i \in c'_{-F}$ prefer r' over r, each of them is thus best off pre-committing to accept r' under full support—contradicting the assumption that

τ was an equilibrium.

Case 2: r is undominated in $\mathcal{R}$. Hence, $r \notin QC(\Gamma)$ implies that $r \not\succ_c \tilde{v}$. Let $c \in C(r)$ denote the proto-coalition that accepts r according to τ, and let (r', c') denote the continuation result in case F proposes (r, c) and it gets rejected. Since (r, c) is not rejected, either $r = r'$ or $r \succ_{c_{-F}} r'$ is satisfied generically. If $r = r'$, then the fact that $r \not\succ_c \tilde{v}$ contradicts the assumption that it is accepted in equilibrium. On the one hand, if there exists $i \in c_{-F}$ such that $r \not\succ_i \tilde{v}$, then this i is better off rejecting (r, c). On the other hand, if $r \not\succ_F \tilde{v}$, then F is best off deviating from proposing r (let r'' denote the continuation outcome; if $r = r''$, then F benefits since r has been delayed, else $r \succ_F r''$ must result, since τ would not be an equilibrium otherwise, and by transitivity $r'' \not\succ_F \tilde{v}$; applied iteratively this contradicts either the finiteness of $\mathcal{R}$ or the fact that F is best off proposing the respective option). Alternatively, if $r \neq r'$, then $r \succ_{c_{-F}} r'$ (as r would not be accepted otherwise), and in this case, the assumption that (r', c') would be the continuation outcome contradicts the arguments made above (note that r' is not dominated in $\mathcal{R}$). $\square$

5 Application to standard examples

Existing applications of proto-coalition bargaining assume the one-round model, where rational F generally choose the option (i.e. the coalition) they prefer most. Deviations from this choice are explained through bounded rationality (i.e. random utility perturbations). One application of the above model of strategic interaction between F and his opponents is that it rationalizes certain deviations from the option F prefers most.

A second application that seems important can be explained best if we recall that subsequent to the formation of the proto-coalition, intra-coalitional negotiations start to allocate cabinet posts (and the like). Most existing studies assume that the proto-coalition chosen by F bargains over the allocation of a cake using a random proposer protocol (an exception is Baron and Diermeier, 2001). A general feature of random proposer models is that a player's expected utility is (weakly) increasing in his recognition probability (i.e. in the probability that he is selected to make the next proposal, see e.g. Eraslan, 2002). Hence, the allocation of proposal power likely becomes a topic during coalition formation. This is neglected in the existing applications of proto-coalition bargaining. For example, Diermeier et al. (2003) assume that the recognition probabilities are exogenous and satisfy (slightly adapting notation and neglecting cases where single parties have absolute majority)

$$\rho_i = \frac{\exp(a_i \lambda \pi_i)}{\sum_{j \in c} \exp(a_j \lambda \pi_j)} \qquad \text{where } \lambda \in \mathbb{R}_+ \text{ and } a_i \in \mathbb{R}_+ \forall i \in c, \tag{13}$$

if c is the proto-coalition selected by F and π_i is the seat share of $i \in c$ in the parliament. Diermeier et al. assume $a_i = 1$ for all i, i.e. symmetry after controlling for seat shares, but this symmetry assumption is likely violated in certain cases. For example, consider the case that i is more important for a successful coalition than j (e.g. because i is ideologically central while j is one of several extreme options). Then $a_i > a_j$ intuitively follows. Naturally, empirical analyses of this hypothesis do not yet exist, but the following theoretical results provide the basis for such an analysis. First, let us define a class of proto-coalition games that endogenize the weights a_i underlying the recognition probabilities in Eq. (13).

Definition 5.1 (Distributive game)**.** Let $\mathbf{A} = \{\mathbf{a} \in \mathbb{N}_0^N \,|\, \sum_{i \in N} a_i = k\}$, for some $k \geq n$, denote the set of allocations and let $C \subset \mathcal{P}(N)$ denote the set of decisive coalitions. The set of feasible proposals is

$$\mathcal{F} = \{(\mathbf{a}, c) \in \mathbf{A} \times C \,|\, \forall i \in c : a_i > 0 \text{ and } \forall j \notin c : a_j = 0\}. \tag{14}$$

For the following analysis, I assume the payoff functions v_i satisfy the following restrictions for all i. (A1) "more is better" in every coalition, (A2) separability of distributive preferences and coalition preferences, (A3) the status quo is bad but better than nothing, and (A4) players prefer small coalitions. In relation to the existing literature, assumptions (A1) and (A4) are implied by typical models of random proposer bargaining, (A3) is standard, and (A2) is satisfied for example under linear separability as assumed by Diermeier et al. (2003). The term $\Delta_i(c, c')$ that is implicitly defined in (A2) represents the compensation i requires when changing from c' to c.

(A1) $a_i > a'_i \Rightarrow v_i(\mathbf{a}, c) > v_i(\mathbf{a}', c)$

(A2) For all c, c' such that $c \cap c' \supseteq \{i\}$ there exists $\Delta_i(c, c') \in \mathbb{Z}$ such that $\forall \mathbf{a}, \mathbf{a}' \in A$: $a_i - a'_i > \Delta_i(c, c') \Leftrightarrow v_i(\mathbf{a}, c) > v_i(\mathbf{a}', c')$

(A3) $a_i > 0 \Rightarrow v_i(\mathbf{a},c) > \tilde{v}_i$ and $a_i = 0 \Rightarrow v_i(\mathbf{a},c) < \tilde{v}_i$

(A4) $i \in c \subset c' \Rightarrow \Delta_i(c,c') < 0$

The first result characterizes the relevant solution sets *QC* and *SG* and shows that *QC* actually equates with the core.

Lemma 5.2. *In distributive games, the quasi-core equates with the core,*

$$QC(\Gamma) = \{(\mathbf{a},c) \in \mathcal{F} \mid \nexists(\mathbf{a}',c') \in \mathcal{F} \text{ such that } (\mathbf{a}',c') \succ_{c'} (\mathbf{a},c)\}, \tag{15}$$

and in addition, $SG = \overline{QC}$ and SG convexifies the core under $\succsim_F$,

$$SG(\Gamma) = \{(\mathbf{a},c) \in \mathcal{F} \mid \exists(\mathbf{a}',c') \in QC(\Gamma) \text{ such that } (\mathbf{a},c) \succsim_F (\mathbf{a}',c')\}. \tag{16}$$

Proof. By (A3), all feasible proposals $(\mathbf{a},c) \in \mathcal{F}$ dominate the status quo, and hence the definition of *QC* simplifies to (15). (A3) also implies that all proposals in $\mathcal{F} \setminus QC$ dominate the status quo, and hence $\overline{QC}$ contains all $(\mathbf{a},c)$ that F weakly prefers to $\min QC$. Finally, by definition of *SG*, (A3) implies that $\min SG$ (under $\succsim_F$) equates with $\min QC$ (under $\succsim_F$). This yields the characterization of *SG*. □

These characterizations of the solution sets will be helpful in the subsequent analysis. The purpose is to show that the solution sets intuitively change as we vary the circumstances of the coalition formation problem. In order to characterize the circumstances, let us next formalize three kinds of restrictions that may be imposed on distributive games. First, the game is *simple* (i.e. a simple majority game) if (i) the complement to any minimal winning coalition (plus F) is also a winning coalition and (ii) all minimal coalitions are of the same size. In the following, $C_{\min} \subset C$ denotes the set of minimal winning coalitions and $|c|$ denotes the cardinality of $c \in C$.

Definition 5.3 (Simple)**.** (i) For all $c \in C_{\min}$ there exists a unique $c' \in C_{\min}$ such that $c \cap c' = \{F\}$. (ii) For all $c, c' \in C_{\min}$, $|c| = |c'|$.

Second, the game is *pure* (i.e. purely distributive) if increases of the own weight a_i lexicographically dominate both one's coalition preferences and one's distributive preferences concerning the allocation a_{-i} between one's opponents. Implicitly, this requires that the compensations defined in (A2) satisfy $\Delta_i(c,c') \in \{0,-1\}$ for all c, c'.

Definition 5.4 (Pure)**.** For all $i \in N$, all $c, c' \in C$, and all $\mathbf{a}, \mathbf{a}' \in \mathbf{A}$: $a_i > a'_i$ implies $v_i(\mathbf{a}, c) > v_i(\mathbf{a}', c')$.

Third, the game is *homogenous* if the players' coalition preferences are aligned in the sense that if one player prefers coalition c to coalition c' (assuming he belongs to both of them), then any other player in both of these coalitions prefers c, too.

Definition 5.5 (Homogenous)**.** For all $\mathbf{a} \in \mathbf{A}$, all $i, j \in N$ and $c, c' \in C$ such that $\{i, j\} \subseteq c \cap c'$: $\Delta_i(c, c') < 0 \Rightarrow \Delta_j(c, c') < 0$.

First, let us look at cases where all three assumptions are satisfied. Then, the outcome most preferred by F uniquely results. In these cases, the solution sets of the infinite horizon games therefore equate with the one of the one-round proto-coalition game that is used in the existing literature.

Lemma 5.6. *If Γ is simple, pure, and homogenous, then $SG(\Gamma) = \{(a^*, c^*)\}$ is a singleton where $c^* \in \arg\max_{c \in C} v_F(\mathbf{a}^*, c)$ and $a_i^* = 1$ for all $i \in c^* \setminus \{F\}$.*

Proof. The proposal $(\mathbf{a}^*, c^*)$ defined above is the unique maximizer of v_F, and hence necessarily in *QC* (thus also in *SG*) if the status quo is bad. It has to be shown that it is unique in *QC*; by the definition of *SG*, $\min QC = \min SG$ under (A3), which then extends the singleton property to *SG*. For contradiction, assume *QC* contains some $(\mathbf{a}, c) \neq (\mathbf{a}^*, c^*)$. Case 1: $c \neq c^*$. By feasibility of $(\mathbf{a}, c)$ and $|c| \geq |c^*|$, an alternative $(\mathbf{a}', c^*) \in \mathcal{F}$ exists such that $a'_i \geq a_i$ for all $i \in c^*$, and by preference homogeneity, all $i \in c^*$ prefer it to $(\mathbf{a}, c)$. Case 2: $\mathbf{a} \neq \mathbf{a}^*$. If $c \neq c^*$ applies as well, then case 1 applies. Otherwise, $c = c^* \in C_{\min}$ applies, i.e. there is $c' \in C_{\min}$ such that $c \cap c' = \{F\}$. Hence $(\mathbf{a}', c') \in \mathcal{F}$ exists such that $a'_i = 1$ for all $i \in c' \setminus \{F\}$, and since $\mathbf{a} \neq \mathbf{a}^*$ implies either $a_i > 1$ for some $i \in c \setminus \{F\}$ or $a_j > 0$ for some $j \notin c$, this additionally allows $a'_F > a_F$. Hence, all $i \in c'$ prefer $(\mathbf{a}', c')$ over $(\mathbf{a}, c)$, the contradiction. $\square$

Now, let us drop "homogeneity" of coalition preferences. In this case, F's potential coalition members may prefer other coalitions than F (e.g. because of their relative ideological positioning). As we will see, this is sufficient to rationalize the phenomenon that F cannot always install his most preferred coalition. Lemma 5.7 first establishes a necessary condition for inclusion in *QC* and *SG*, namely that every solution must sustain a minimal winning coalition and minimal weights for F's opponents, and second it shows by example that all but one minimal winning coalition may be included in SG. In contrast, recall that if all players' coalition preferences are aligned ("homogenous") then only F's favorite coalition is included in SG.

Lemma 5.7. *Assume Γ is simple and pure. (i) All $(\mathbf{a}^*, c^*) \in SG$ satisfy $c^* \in C_{\min}$ and $a_i^* = 1$ for all $i \in c^* \setminus \{F\}$. (ii) There are constellations of preferences where all but one $c \in C_{\min}$ are sustained in SG.*

Proof. Point (i): First, we show $c^* \in C_{\min}$ for all $(\mathbf{a}^*, c^*) \in QC$. For contradiction, assume there exists $(\mathbf{a}', c') \in QC$ such that $c' \notin C_{\min}$. Thus, there is $c \in C_{\min}$ such that $c \subset c'$, and by (A4) this implies $\Delta_i(c, c') < 0$ for all $i \in c$. Define $\mathbf{a}$ such that $(\mathbf{a}, c) \in \mathcal{F}$ and $a_i = a_i'$ for all $i \in c \setminus \{F\}$. By $\Delta_i(c, c') < 0$, $v_i(\mathbf{a}, c) > v_i(\mathbf{a}', c')$ for all $i \in c$ results, the contradiction. Knowing $c^* \in C_{\min}$ for all $(\mathbf{a}^*, c^*) \in QC$, we next show $a_i^* = 1$ for all $i \in c \setminus \{F\}$. Assume there exists $(\mathbf{a}', c) \in QC$ such that $c \in C_{\min}$ but $a_i > 1$ for some $i \neq F$. By simplicity, there exists $c'' \in C_{\min}$ such that $c'' \cap c' = \{F\}$, and hence there also exists $(\mathbf{a}'', c'') \in \mathcal{F}$ such that $a_i'' = 1$ (implying $a_i'' > a_i'$) for all $i \in c'' \setminus \{F\}$. Now, $|c''| = |c'|$ implies $a_F'' > a_F'$, and hence $v_i(\mathbf{a}'', c'') > v_i(\mathbf{a}', c')$ for all $i \in c''$ follows (the contradiction). Thus the result is established for all $(\mathbf{a}^*, c^*) \in QC$, and since $|c| = |c'|$ for all $c, c' \in C_{\min}$ under simplicity, this implies $a_F^* = k - |c^*|$ for all $(\mathbf{a}^*, c^*) \in QC$. Hence, $a_F^* > a_F$ for all $(\mathbf{a}, c) \in \mathcal{F}$ where $c \notin C_{\min}$, and by purity, this implies $v_F(\mathbf{a}^*, c^*) > v_F(\mathbf{a}, c)$. Due to $\min QC = \min SG$ under $\succsim_F$, this extends the claim to all $(\mathbf{a}^*, c^*) \in SG$. *Point (ii):* Let $a_{F,\max}$ denote the maximal bargaining weight a_i that F can attain subject to feasibility $(\mathbf{a}, c) \in \mathcal{F}$, and define the worst-case coalition $c \in \arg\min_{c' \in C_{\min}} v_F(\mathbf{a}, c')$ subject to $a_F = a_{F,\max}$. Let $\overline{c}$ denote the complement to c, i.e. the unique $\overline{c} \in C_{\min}$ such that $c \cap \overline{c} = \{F\}$. Construct preferences such that $\overline{c}$ is the second-worst coalition in F's eyes, i.e. $\Delta_F(\overline{c}, c') = 0$ for all $c' \neq c$, and such that all $i \neq F$ have coalition preferences that oppose those of F, i.e. for all $c, c' \in C_{\min}$ such that $c \cap c' \ni i$: $\Delta_i(c, c') = -1 - \Delta_F(c, c')$. Due to $\Delta_F(\overline{c}, c) = -1$, i.e. F prefers $\overline{c}$ to its (unique) complement, there must be some $\mathbf{a}$ such that $(\mathbf{a}, \overline{c}) \in QC$ (by construction, for alternative c', there is at least one $i \neq F$ that prefers $\overline{c}$ to c'). By convexity of SG under $\succsim_F$ and the fact that $\overline{c}$ is the second-worst coalition in F's eyes, this implies that all $c' \in C_{\min}$ but c are sustained in SG. $\square$

An alternative explanation for the possibility that F is unable to install his favored coalition is that the majority game is not simple, i.e. that the complement to a decisive coalition plus F is not decisive. To illustrate this, let us look at the case that the formateur has no voting power himself. If the majority game is simple otherwise, this implies that the intersection of all pairs of minimal winning coalitions has two elements, F and one other player. The next result shows that in such cases, *QC* sustains only F's favorite coalition, while *SG* sustains all minimal winning coalitions. Thus, if F's opponents cannot make positive pre-commitments, any minimal winning coalition may result in equilibrium.

Lemma 5.8. *Assume Γ is pure, homogenous, and satisfies $|c| = |c'|$ as well as $|c \cap c'| = 2$ for all $c, c' \in C_{\min}$. Then, the following applies.*

1. $(\mathbf{a}^*, c^*) \in QC$ *only if* $c^* \in \arg\max_{c \in C} v_F(\mathbf{a}^*, c)$ *and* $a_i^* \in \{1, 2\}$ *for all* $i \neq F$
2. *There is at least one* $(\mathbf{a}^*, c^*) \in QC$ *where* $a_i^* = 2$ *for some* $i \neq F$
3. *For all* $c \in C_{\min}$*, there exists some* $\mathbf{a}$ *such that* $(\mathbf{a}, c) \in SG$

Proof. Point 1: $c \in C_{\min}$ follows from (A4). Given this, $c = c^*$ follows from homogeneity and $|c| = |c^*|$ for all $c \in C_{\min}$. For any alternative $(\mathbf{a}, c)$, there exists $(\mathbf{a}', c^*)$ that F prefers by definition of c^* and that all $i \in c^* \cap c$ prefer by homogeneity. Finally, assume (for contradiction) that there exists $(\mathbf{a}^*, c^*)$ such that $a_i^* > 2$ for some $i \neq F$. Take any $c \in C_{\min}$ such that $c \cap c^* \ni j \neq i$ and define $\mathbf{a}$ such that $a_j = a_j^* + 1$ and $a_k =$ for all $k \notin \{j, F\}$. This implies $a_F > a_F^*$ and hence $v_k(\mathbf{a}, c) > v_k(\mathbf{a}^*, c^*)$ for all $k \in c$, which contradicts $(\mathbf{a}^*, c^*) \in QC$. *Point 2:* Define $a^* \in \arg\max_{\mathbf{a}} v_F(\mathbf{a}, c^*)$ s.t. $a_i = 2$ for at least one $i \neq F$. I claim $(\mathbf{a}^*, c^*) \in QC$. Define $i \neq F$ as the player with $a_i^* = 2$. By purity, i is unique. Hence, for all $(\mathbf{a}, c)$ such that $v_F(\mathbf{a}, c) > v_F(\mathbf{a}^*, c^*)$, $a_j = 1$ for all $j \neq F$ holds true. Also for all such $(\mathbf{a}, c)$, $c \cap c^*$ contains at least one $j \neq F$, and by homogeneity and $a_j \leq a_j^*$, $v_j(\mathbf{a}, c) < v_j(\mathbf{a}^*, c^*)$. Hence, no such $(\mathbf{a}, c)$ dominates $(\mathbf{a}^*, c^*)$. *Point 3:* Since $(\mathbf{a}^*, c^*)$ as constructed in point 2 is in *QC*, and since *SG* convexifies *QC* under $\succsim_F$, all $(\mathbf{a}, c)$ satisfying $c \in C_{\min}$ and $a_i = 1$ for all $i \neq F$ are in *SG*. $\square$

Finally, by relaxing "purity" we can investigate the case that F strongly favors a specific coalition. Intuitively one may suspect that the members of F's favorite coalition are able to exploit F's desire to coalesce with them and extract additional bargaining power. Again, structural empirical analyses of this intuition do not yet exist, but as I show next, this intuition is indeed compatible with the solution sets derived above. Formally, define the "relative (opportunity) costs" $RC(c) = \min_{c'} \Delta_F(c,c')$ as the compensation required by F for choosing c instead of the best alternative to c. The unique minimizer of $RC(c)$ is the coalition favored by F, i.e. it is

$$c^* \in \arg\max_{c \in C} \max_{\mathbf{a} \in \mathbf{A}} v_F(\mathbf{a},c). \tag{17}$$

The following lemma shows that F's favored coalition c^* necessarily results, while its members can extract up to $RC(c^*)$ units of bargaining weight from F during coalition formation. The resulting allocation of this bargaining weight is not restricted in equilibrium (i.e. any of c^*'s members may get it). For simplicity, the following assumes that F prefers c^* equally strongly over all alternatives to c^* in terms of $\Delta_F(c^*,\cdot)$.

Lemma 5.9. *Assume Γ is simple, homogenous, and $\Delta_F(c^*,c') = \Delta_F(c^*,c'')$ for all c',c'' such that $c' \neq c^*$, $c'' \neq c^*$, and $c^* \in \arg\min_c RC(c)$. Then, both QC and SG are the sets of all $(\mathbf{a}^*,c^*) \in \mathcal{F}$ satisfying*

$$c^* \in \arg\max_{c \in C} \max_{\mathbf{a} \in \mathbf{A}} v_F(\mathbf{a},c) \qquad\qquad |c^*| \leq \sum_{i \neq F} a_i^* < |c^*| - RC(c^*). \tag{18}$$

Proof. All $(\mathbf{a}^*,c^*)$ satisfying the above constraints also satisfy $v_F(\mathbf{a}^*,c^*) > v_F(\mathbf{a},c)$ for all $(\mathbf{a},c) \in \mathcal{F}$ where $c \neq c^*$ (by definition of RC). Hence, all such $(\mathbf{a},c) \notin QC$ in general, and $(\mathbf{a}^*,c^*) \notin QC$ only if there exists $(\mathbf{a},c^*) \in \mathcal{F}$ such that $\mathbf{a}$ weakly Pareto dominates $\mathbf{a}^*$. The latter applies only if $\sum_{i \in c^*} a_i > \sum_{i \in c^*} a_i^*$ (over all $i \in c^*$), which in turn violates feasibility. Hence, all these $(\mathbf{a}^*,c^*)$ are in QC and hence in SG. It remains to show that no $(\mathbf{a}^*,c^*)$ is in QC that satisfies $\sum_{i \neq F} a_i^* \geq |c^*| - RC(c^*)$. All such $\mathbf{a}^*$ thus satisfy $a_F^* \leq k - |c^*| + RC(c^*)$. Define $c \in C$ such that $c \cap c^* = \{F\}$. By assumption $RC(c^*) = \Delta_F(c^*,c)$, and thus by (A2), $v_F(\mathbf{a}',c) > v_F(\mathbf{a}^*,c^*)$ if $a_i^* - a_i' \leq \Delta_F(c^*,c)$, i.e. if $a_F = k - |c| + 1$. The latter applies if $a_i = 1$ for all $i \in c \setminus \{F\}$, and is therefore feasible. Hence, no $(\mathbf{a}^*,c^*)$ is in QC where $a_F^* \leq k - |c^*| + RC(c^*)$. □

6 Conclusion

The paper analyzed the formation of proto-coalitions without the assumption of a finite time horizon. The existing literature focuses on one-round models, which imply that a rational F forms his favored proto-coalition if the payoffs under the caretaker government are not prohibitive (and they often are not, since the caretaker has limited legitimacy). The one-round assumption prevents the possibility that F's potential coalition partners strategically interact with F, say to implement their favored coalitions or to extract bargaining power from F. I have characterized the solution sets of the infinite horizon games (assuming semi-stationary strategies) for two alternative regimes. If F's opponents can publicly pre-commit, then the "quasi core" is implemented, which in turn equates with the core if the status quo is "bad." If F's opponents cannot pre-commit, then the solution can be characterized as a self-generating set, which exactly convexifies the core if the status quo is bad. In this sense, we can conclude that proto-coalition bargaining implements or convexifies the core.

These characterizations of the solution sets are straightforwardly applicable in empirical analyses if the valuations of all participants are well defined (as in Diermeier et al., 2002, 2003, for example). This seems promising, since the infinite horizon game has been shown (in Section 5) to be compatible with the phenomena that F's favored coalition may not form and that F may have to give up bargaining power toward "strong" coalition partners. In the existing literature, the former has not been rationalized yet for non-prohibitive default payoffs, and the latter has not yet been formalized explicitly. These extensions of the proto-coalition model should facilitate future empirical work.

Finally, let me comment on the fact that the characterization of the equilibrium outcomes yielded a solution set, but obviously not a probability distribution over the solution set. There does not seem to be a generally accepted approach to the resolution of such a multiplicity of equilibria for empirical analyses. However, empirical applications of the one-round game assumed logit choice of F, rather than rational choice, to account for deviations from F's favored choice. The same approach can be applied in applications of the infinite horizon game if we substitute "logit equilibrium" for logit choice (following McKelvey and Palfrey, 1995, in general, and

Breitmoser et al., 2010, for dynamic games)—and in the case of logit equilibria, the standard approach is to focus on the equilibria located on the "the principal branch" (Turocy, 2005, 2010). In this way, the multiplicity is resolved immediately, while the interaction between F and the other parties can be modeled accurately.

References

Austen-Smith, D. (2000). Redistributing income under proportional representation. *Journal of Political Economy*, 108(6):1235–1269.

Banks, J. and Duggan, J. (2006). A general bargaining model of legislative policy-making. *Quarterly Journal of Political Science*, 1:49–85.

Baron, D. and Diermeier, D. (2001). Elections, governments, and parliaments in proportional representation systems. *The Quarterly Journal of Economics*, 116(3):933–967.

Baron, D. and Ferejohn, J. (1989). Bargaining in legislatures. *The American Political Science Review*, 83(4):1181–1206.

Breitmoser, Y., Tan, J., and Zizzo, D. (2010). Understanding perpetual r&d races. *Economic Theory*, 44(3):445–467.

Cardona-Coll, D. and Mancera, F. M. (2000). Demand bargaining in legislatures. *Economic Theory*, 16(1):163–180.

Diermeier, D. (2006). Coalition government. *Oxford handbook of political economy*, pages 162–179.

Diermeier, D., Eraslan, H., and Merlo, A. (2002). Coalition governments and comparative constitutional design. *European Economic Review*, 46(4-5):893–907.

Diermeier, D., Eraslan, H., and Merlo, A. (2003). A structural model of government formation. *Econometrica*, 71(1):27–70.

Diermeier, D., Eraslan, H., and Merlo, A. (2007). Bicameralism and government formation. *Quarterly Journal of Political Science*, 2:227–252.

Diermeier, D. and Merlo, A. (2000). Government turnover in parliamentary democracies. *Journal of Economic Theory*, 94(1):46–79.

Eraslan, H. (2002). Uniqueness of stationary equilibrium payoffs in the Baron-Ferejohn model. *Journal of Economic Theory*, 103(1):11–30.

Kalandrakis, T. (2006). Regularity of pure strategy equilibrium points in a class of bargaining games. *Economic Theory*, 28(2):309–329.

Kultti, K. (2000). A model of random matching and price formation. *European economic review*, 44(10):1841–1856.

McKelvey, R. and Palfrey, T. (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10(1):6–38.

Merlo, A. and Wilson, C. (1995). A stochastic model of sequential bargaining with complete information. *Econometrica*, 63(2):371–399.

Merlo, A. and Wilson, C. (1998). Efficient delays in a stochastic model of bargaining. *Economic Theory*, 11(1):39–55.

Persson, T. (2002). Do political institutions shape economic policy? *Econometrica*, 70(3):883–905.

Persson, T. and Tabellini, G. (2004). Constitutional rules and fiscal policy outcomes. *American Economic Review*, 94(1):25–45.

Rubinstein, A. and Wolinsky, A. (1985). Equilibrium in a market with sequential bargaining. *Econometrica*, 53(5):1133–1150.

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

Shapley, L. and Shubik, M. (1966). Quasi-cores in a monetary economy with non-convex preferences. *Econometrica*, 34(4):805–827.

Shimomura, K. (1997). Quasi-cores in bargaining sets. *International Journal of Game Theory*, 26(3):283–302.

Turocy, T. (2005). A dynamic homotopy interpretation of the logistic quantal response equilibrium correspondence. *Games and Economic Behavior*, 51(2):243–263.

Turocy, T. (2010). Computing sequential equilibria using agent quantal response equilibria. *Economic Theory*, 42(1):255–269.

Winter, E. (1994). The demand commitment bargaining and snowballing cooperation. *Economic Theory*, 4(2):255–273.

Wooders, J. (1998). Matching and bargaining models of markets: approximating small markets by large markets. *Economic Theory*, 11(1):215–224.