# Stable Sequences of Political Coalitions

Stefan B. Imhof [*]

Department of Economics

European University Institute, Florence

email: stefan.imhof@iue.it

April 2003

**Abstract**

This paper explores a sequential coalition formation game among political parties. We introduce the non-cooperative concept of stable sequences of coalitions, a general solution to sequential coalition formation games. The main results are i) the order of the agenda matters for the equilibrium outcome, ii) punishment strategies can support otherwise unstable coalition structures, in particular the phenomenon of "strange bedfellows" can arise in the first round, and iii) a party which is median in all decisions is always better off in the sequential game than in a single coalition formation over two decisions, while the converse is not true.

JEL classification numbers: C72, D71, D72

Keywords: coalition formation, legislation

[*] I thank Karl Schlag and seminar participants at the European University Institute for helpful comments.

# 1 Introduction

This paper aims to contribute to the ongoing debate on the formation of political coalitions in parliamentary democracies with electoral laws that are based upon the principle of (not necessarily perfectly) proportional representation, found e.g. in many continental European countries.[1]

In particular, we show that coalitional structures, which seem to be unstable, can indeed be stable if parties are allowed to form coalitions not only once and for all, but to form coalitions over each single political issue sequentially. The key is that later negotiations can be made contingent on the outcomes of earlier coalitions.

Since proportional representation does not usually generate absolute majorities, governments in these systems typically consist of coalitions that are formed by different political parties. In most cases, these coalitions provide the government with an absolute majority in the main legislative body. Arguably, this coalition formation process is inherent to parliamentary democracies and plays an important role in the political process.

Once a coalition is formed successfully, an agenda is formulated, and government proposals, which are submitted to parliament are approved by the ruling parties without further amendment. Government proposals in parliamentary democracies outweigh other forms of parliamentary proposals. The same a fortiori holds true for proposals being approved by the parliament.

In many cases, a coalition is formed by parties, which perceive each other as ideologically close. This confirms traditional game theoretical models of political coalition formation, in which so-called connected coalitions are not only the most likely, but also the most stable (lasting) coalitions. In these models it is assumed that parties can be ideologically ordered on a one-dimensional policy space. The term connectedness refers to coalitions in which all members take on positions, which are adjacent to each other. The classical reference for connectedness of political coalitions is Axelrod [3], who argues that the reason for this is that political parties are not only office-seeking, as earlier theories assumed (von Neumann and Morgenstern [16], Riker [17]), but also try to implement policies, which are as close as possible to their ideal position.[2] All

[1] Recently, some European countries, especially France and Italy, reformed their electoral laws towards a majority votes system, which is more likely to generate absolute, and hence stable, majorities. A typical example of a parliamentary democracy with a majority votes system is the UK, in which the simple majority of votes in a constituency suffices to win a seat.

[2] Using a cardinal measure on a single-dimensional policy space, De Swan [13] develops a theory of closed minimal range coalitions in order to select among all possible connected minimal winning coalitions.

of these theories predict the formation of minimal winning coalitions, i.e., coalitions, in which each party is decisive for the coalition in order that it has an absolute majority (or some other necessary quorum) of votes. However, while traditional theories successfully predict in many cases, we frequently observe coalition structures, which these theories cannot explain.

The first one is the phenomenon of "strange bedfellows", where ideologically remote parties form a coalition. Brams et al. [9] show that such disconnected coalitions can be the result of different perceptions of the distance between the ideal positions of the parties. This might happen if players have single-peaked preferences over a single-dimensional policy space, in which ideal positions can be ordered ordinally, but not cardinally. However, once positions can be ordered cardinally, i.e., all parties have the same perception of the location of all other parties, again only connected coalitions will be formed.

The second "non-intuitive" coalition structure, which is observed especially in local legislations, but also in minority governments, is a seemingly unstable one, in which one party forms coalitions over single decisions with different coalition partners.[3]

More recent work on political coalition formation challenges two fundamental assumptions, which are widely used in the literature. These assumption are i) a single-dimensional policy space, and ii) the one-shot character of the coalition formation process. As we show below, they are closely related and can, in a certain sense, be considered equivalent.

### Outline and main results

This paper aims to provide an insight into the relation between dimensionality and sequentiality of the coalition formation problem in a model of a parliament, which consists of three parties. The assumption of only three parties facilitates the analysis in the sense that any non-degenerate coalition is a winning coalition, given that an absolute majority is the necessary quorum for deciding upon a policy issue. The bargaining solutions to each coalition depend on the ideological distance and the ideal policy of the coalition members, which are known to the parties, so that the game reduces to a hedonic coalition formation game. We analyze stable coalition structures in three different cases. We start with the case in which parties form a coalition over a single decision, where decision means that parties are located on a one-dimensional policy space. Then we assume that parties form a single coalition over two decisions,

[3]Especially in local governments, less controversial decisions are often made by the grand coalition.

i.e. a single coalition in a two-dimensional policy space without renegotiation, a coalition formation process, which is substantially similar to an enforceable coalition contract. In the main model, parties sequentially form coalitions over two decisions, i.e. the coalition formation is repeated, while the location of the parties is not assumed to be fixed over time. This reflects the observation that political parties can take on different relative positions on different policy issues, such that a party can be median in one particular issue, but not in another. We always assume perfect and complete information.

In order to solve the game, we define a non-cooperative concept of stability, which combines the type of coalitional deviations as found in the concept of perfectly strong Nash equilibrium (Aumann [2]) with a self-enforcement requirement similar to perfectly coalition-proof Nash equilibrium (Bernheim, Peleg, Whinston [6]). It turns out that our solution can be interpreted as renegotiation-proof Nash equilibrium (Xue [18]), in which players are forward-looking with respect to their coalitional behavior. In contrast to the latter, our solution can be completely defined in terms of coalitional deviations only, which also induces a clear relation between stable coalition structures and the cooperative solution of the core of a non-transferable utility (NTU) game.

The main results of the analysis can be summarized as follows. In the single decision case we find that there is always a stable minimal winning coalition. In particular, we confirm previous results that there are no disconnected stable coalitions, implying also that there is always at least one non-degenerate coalition which is never stable, irrespective of the allocation of bargaining powers. Moreover, in contrast to previous authors (e.g. Kirchsteiger and Puppe [15]), and as a result of a different approach to the bargaining solution, we find that the grand coalition can be stable, too. Generally, for a given allocation of bargaining powers, a coalition is more likely to be stable in the single decision case if the ideological distance between the coalition members is smaller.

In the case of a single coalition formation over two decisions, all non-degenerate coalitions can be stable for some allocation of bargaining power. Yet, an ideologically remote coalition can only be stable, if all ideologically closer coalitions are stable for some (different) allocation of bargaining powers. In particular, if a party is the median party on both policy issues, then a coalition of the two other parties is never stable, i.e. the result of the single decision problem is re-established.

The sequential coalition formation game differs substantially from both single coalition formation games. Clearly, the nature of the problem allows for changing coalitions over time, and it turns out that any sequence of stable coalitions will also be an equilibrium of the dynamic game. In contrast to the single coalition formation over two decisions, there is always a sequence of coalitions, which consists of the same two parties in both periods, and which is never stable for any allocation of bargaining power.

Moreover, it is shown that the outcomes of the two different regimes cannot be Pareto ranked. Only a party, which is median in both decisions, is always better off in the sequential coalition formation game. This also implies that the choice of a coalition formation regime is relevant to the individual outcome of the coalition formation process.[4]

Another important result, which contrasts sharply with a single decision problem, is the existence of disconnected coalitions in the first stage of the sequential game. Intuitively, this means that if political parties are aware of the fact that they will have to repeat the coalition formation game under potentially different conditions, they may be willing to form coalitions with ideologically remote partners. The "strange bedfellow" phenomenon can be supported in equilibrium only if there are multiple stable coalitions in the second round. In this sense, we provide an alternative explanation to Brams et al. [9].

The paper is organized as follows. Section 2 presents the model. We then introduce the concept of stable sequences of coalition in section 3. Section 4 applies the solution concept to the different coalition formation games. The analysis follows the same order as the presentation above. The main results are compared in section 5. Section 6 concludes.

# 2 The Environment

The parliament consists of three parties $i \in \{1, 2, 3\}$ with $p_i \leq \frac{1}{2}$ for all $i$ relative number of seats, i.e., $\sum_i p_i = 1$. In order to make a decision, parties have to form coalitions, since a decision can be taken if and only if it is approved by an absolute majority of votes. Denote by $\mathbf{J}$ the set of all subsets of the set $\{1, 2, 3\}$, i.e., the set of all coalitions including degenerate coalitions consisting of a single party, by $J$ a typical element of $\mathbf{J}$, and let $\mathbf{J}^i = \{J \in \mathbf{J} : i \in J\}$ be the set of coalitions containing party $i$. Analogously, define $\mathbf{J}^{-i} = \{J \in \mathbf{J} : i \notin J\}$ as the set of coalitions, which do not contain party $i$. Hence, $\mathbf{J} = \{\mathbf{J}^i, \mathbf{J}^{-i}\}$. In the legislative session two decisions, $t = 1, 2$, are to be made. For each single decision, parties have a most preferred pol-

[4]The timing of the decisions is assumed to be exogenous. However, the timing clearly plays an important role in the coalitions formation process. If the timing were endogenous, "winners" and "losers" might change places.

icy, which is assumed to be located on the half-open interval $[0, \infty)$. These positions are not required to be fixed over the sequence of decisions. A position of a party in a decision will be denoted by $x_{it}$, with $i = 1, 2, 3$ and $t = 1, 2$. Recall that by the assumption of perfect information these positions are known to all parties.
If two or more parties agree on forming a coalition, they bargain over the set of alternatives and for each decision they choose some alternative $B(J_t) \in [0, \infty)$, where $J_t = \{i : i \in J_t\}$ is a coalition at time $t$. $\mathbf{J}$ can be written as $\mathbf{J} = \{1, 2, 3, 12, 13, 23, 123\}$, $\mathbf{J}^1 = \{1, 12, 13, 123\}$, $\mathbf{J}^{-1} = \{2, 3, 23\}$, and so on. We write a sequence of coalitions as $J = (J_1, J_2)$ with $J \in \mathbf{J} \times \mathbf{J}$. Whenever unequivocal, in particular when only a single decision is considered, we omit the subscript associated with $t$.
A winning coalition is defined as coalition $J \in \mathbf{J}$ with $\sum_{i \in J} p_i > \frac{1}{2}$. $p_i \leq \frac{1}{2}$ implies that all non-degenerate coalitions in $\mathbf{J}$ are winning coalitions, and that, in particular, all coalitions consisting of two parties are minimal winning coalitions. A coalition $J \in \mathbf{J}$ is a minimal winning coalition if and only if for all $i \in J$, $\sum_{j \in J} p_j - p_i \leq \frac{1}{2}$. [5]

**Preferences**

Parties are assumed to have preferences over coalitions and decision outcomes. Office seeking behavior is reflected by the assumption that parties, regardless of the policy outcome, always strictly prefer to be part of a decision-making coalition than any other coalition. That is, parties only care about policy outcomes if they are members of a coalition. In this sense, preferences are partially lexicographic: while preferences are lexicographic over coalitions, they are not over policy outcomes, since parties are indifferent between two policy decisions if they are not part of the coalition, which makes the decision. Once in a coalition, the policy outcome depends on the parties' weight in the coalition, $\frac{p_i}{\sum_{j \in J} p_j}$, and the distance between the coalition member's ideal positions.
Formally, consider first preferences over a single decision problem, and let $(J, B(J)) \in \mathbf{J} \times R_+$. Then $(J, B(J)) \succ_i (J', B(J'))$ if and only if either $J \in \mathbf{J}_i$ and $J' \in \mathbf{J}_{-i}$ or if $J, J' \in \mathbf{J}_i$ and $|B(J) - x_i| < |B(J') - x_i|$. Moreover, $(J, B(J)) \sim_i (J', B(J'))$ if and only if either $J, J' \in \mathbf{J}_{-i}$ for all $B(J), B(J') \in [0, \infty)$ or $J, J' \in \mathbf{J}_i$ and $|B(J) - x_i| = |B(J') - x_i|$. This implies that - once a party is part of a coalition - preferences over decision outcomes are single-peaked and symmetric on the interval $[0, \infty)$.
Now let $(J, B(J)) \in \mathbf{J}^2 \times R_+^2$, i.e., there are two decisions to make and $(J, B(J))$

[5] Clearly, this also implies that the Shapley value is the same for all parties, regardless of their relative size. Let $V = [q; p_1, p_2, p_3]$ with $q = 1/2$. Since $p_i < 1/2$ for all i, we have $v(i) = 0$, and $v(12) = v(13) = v(23) = v(123) = 1$. Then the Shapley value is given by $f_i(v) = 1/3$, for all i.

now is a pair of vectors of length 2. Then $(J, B(J)) \succ_i (J', B(J'))$ if and only if either $|\{J_t : i \in J_t\}| > |\{J'_t : i \in J'_t\}|$ or if $|\{J_t : i \in J_t\}| = |\{J'_t : i \in J'_t\}|$ and $\sum_{t|i\in J_t} |B(J) - x_{it}| < \sum_{t|i\in J_t} |B(J') - x_{it}|$. Moreover, $(J, B(J)) \sim_i (J', B(J'))$ if and only if either $J, J' \in \mathbf{J}^2_{-i}$, or $|\{J : J \in \mathbf{J}^2_i\}| = |\{J' : J' \in \mathbf{J}^2_i|$ and $\sum_t |B(J) - x_{it}| = \sum_t |B(J') - x_{it}|$. This implies that utility is additive over sequential decisions, and that parties do not discount utility over the legislative session.
Hence, a party is fully described by the triple $(\succeq_i, x_i, p_i)$.

### The bargaining solution

Once a coalition is formed, its members are assumed to agree on a policy according to a given rule. In particular, they are assumed to implement an average of their ideal policies weighted by their relative power within the respective coalition.[6]
The bargaining solution is then given by

$$B(J_t) = x_{it} + \sum_{j\in J_t} \frac{p_j}{\sum_{k\in J_t} p_k}(x_{jt} - x_{it})$$

Subsequently, for notational convenience, we will use the distance function $d_t(ij) = |x_{jt} - x_{it}|$ such that $d_t(ji) = d_t(ji)$, and $-d_t(ij)$ whenever unequivocal.

# 3 Solution concept

A single decision problem consists of a coalition formation and a bargaining solution. Coalitions are assumed to be formed in a non-cooperative way. The solution to the bargaining problem, i.e., the decision outcome, is given by $B(J)$. Thus, a single decision can be seen as a reduced two-stage game of coalition formation, and a strategy is an action $s_i \in \mathbf{J}^i$. In the case of two decisions, a strategy of party $i$ is a complete contingent plan of actions in each of the decisions, and a single strategy therefore can be written as $s_i = (s_{i1}, s_{i2}(J_1))$. We write as $S_i$ the strategy space of party $i$. $S = \times_i S_i$, and a strategy vector is denoted as $\mathbf{s} = (s_1, s_2, s_3)$. A coalition is formed if and only if there is a $J \in \mathbf{J}$ such that $s_i = s_j$ for all $i, j \in J$.
The solution we look for is a stable coalition, or (in the case of multiple decisions)a stable sequence of coalitions over a sequence of decisions. Loosely speaking, for a

[6]This solution was proposed by Alesina and Rosenthal [1] for a two-party system. Essentially, the solution is the outcome of a war of attrition with an exogenously given winning probability. See, e.g., Bulow and Klemperer [10] for applications to problems in industrial organization, or Carmignani [11] and [12]. The relative power within a coalition can also be interpreted as the probability that a party can implement its ideal policy. $B(J)$ is then the expected outcome of the bargaining game

sequence of coalitions to be stable we require that there is no other sequence of coalitions which is preferred by all members of a deviating coalition, given the strategy vector which induces the stable sequence, and which is itself immune to mutually beneficial deviations of any other coalition. In other words, a sequence of coalitions is stable if it is immune to self-enforcing deviations by any other sequence of coalitions. Note that while we allow for deviations of any subset of parties, and not only for those of proper subsets of coalitions, we do not allow for any kind of deviation, except for self-enforcing ones. These requirements are intuitive, since if on the one hand we require coalitions to be immune to joint deviations, the deviating coalition should be subject to the same restriction. On the other hand, if parties can freely communicate before forming a coalition, we should not only allow for proper subsets of coalitions to deviate in such a self-enforcing way, but also any subset of players.

Obviously, in a non-cooperative setting this requires a stronger equilibrium concept than perfect Nash equilibrium. Indeed, we take an intermediate position between two solution concepts, which are widely described in the literature, the concepts of (perfectly) strong Nash equilibrium ($SNE$) by Aumann [2], and (perfectly) coalition-proof Nash equilibrium ($CPNE$) by Bernheim, Peleg, and Whinston [6]. A Nash equilibrium is strong if and only if no coalition, taking the action of its complement as given, can agree upon a mutually profitable deviation. That is, strong Nash equilibrium allows for deviations by any conceivable coalition. In turn, an equilibrium is said to be coalition-proof if and only if it is Pareto efficient within the class of self-enforcing agreements, where self-enforcing means that no proper subset of a coalition, taking the actions of its complement as fixed, can deviate in a mutually profitable way.

Our concept of stability is weaker than $SNE$, since it does not allow for any coalitional deviations, but only for self-enforcing ones. Hence, all strong Nash equilibria will generate stable coalitions. There is no immediate unambiguous inclusive relation between stability and $CNPE$ (recall that all $CNPE$ are $SNE$), because on the one hand self-enforceability is not restricted to deviations of proper subsets of deviating coalitions, which makes the concept of stability a priori more demanding than $CNPE$. On the other hand, for the same reason we do not allow for deviations, which induce further deviations of other coalitions than subsets of deviating coalitions, so that stability is weaker than $CNPE$. Since this clearly holds for first deviations as well, it turns out that all stable coalitions are are induced by some (perfectly) coalition-proof strategy vector, while the reverse is not true.

It turns out that our requirements for stability relate closely to the refinement concept of (re)negotiation-proof Nash equilibria ($NPNE$), which was introduced into by Xue [18] in order to improve upon the nestedness restriction of $CNPE$, without being as restrictive as $SNE$. It captures the idea that players can freely suggest and object to coalitional deviations before (each stage of) a game is played, thereby avoiding myopic deviations that can happen in $CPNE$, since players anticipate any possible further deviation of any other coalition.

In the following paragraphs we define the concept of stability in a more rigorous way. While it turns out that in a single decision problem, $SNE$, $CPNE$, and stability are equivalent, this result does not hold true for the dynamic game of two decisions. $SNE$ as well as $CPNE$ require the solution to be efficient with respect to some set of agreements (the former with respect to the entire feasible payoff space of the underlying game, the latter with respect to self-enforcing agreement in the sense above). In the present model there is a straightforward relation between efficiency and the set of coalitions. Lemma 1 shows that all non-degenerate coalitions are efficient and vice versa.

**Lemma 1** *A coalition $J \in \mathbf{J}$ is efficient if and only if it is non-degenerate.*

Proof: Necessity follows from the assumptions that $p_i < 1/2$ and that a decision can only be made with an absolute majority. Since parties always (weakly) prefer a decision to be made rather than no decision at all, all degenerate coalitions are Pareto dominated by any non-degenerate one.

For sufficiency note that no coalition $(ij)$ can Pareto dominate another coalition with two members, since for all $i$, $J_{i\in J} \succ_i J_{i\notin J}$, which leaves the grand coalition as only candidate for a Pareto superior solution. However, by the fact that $B(12) < B(123) < B(23)$, and generically $B(13) \neq B(123)$, either 1 or 3 or both are worse off in (123) than in any other non-degenerate coalition. $\|$

For the following definition of stability note that any sequence of coalitions $J$ is induced by some strategy vector $\mathbf{s} \in S$. Clearly, the associated strategy vector is not necessarily unique.

**Definition 1** *Let $\Psi(\mathbf{s}) = J$, i.e. the mapping which associates the strategy vector $\mathbf{s}$ with its induced sequence of coalitions $J$. Then $J'$ is a deviation by players in $M \subset \{1, 2, 3\}$, denoted by $J \to_M J'$, if there are $\mathbf{s}, \mathbf{s}'$ such that $\Psi(\mathbf{s}) = J$, $\Psi(\mathbf{s}') = J'$ with $s'_i = s_i$ for all $i \notin M$.*

*i) A deviation $J \to_M J'$ is self-enforcing if $J' \succeq_i J$ for all $i \in M$, and $J' \succ_j J$ for some $j \in M$.*

*ii) A deviation $J \rightarrow_M J'$ is strongly self-enforcing, if it is self-enforcing, and if there are no $J'', M'$ such that $J' \rightarrow_{M'} J''$ is self-enforcing.*

**Definition 2** *(Stability)*
*i) $J$ is said to be not stable with respect to $J'$, if there is a $M$ such that $J \rightarrow_M J'$ is strongly self-enforcing.*
*ii) A sequence of coalitions $J = (J_1, J_2)$ is stable if there is no strongly self-enforcing deviation $J \rightarrow_M J'$.*
*iii) In a single decision, a coalition $J_t$ is period-stable if there is no $M \subset \{1, 2, 3\}$, such that a deviation $J_t \rightarrow_M J'_t$ is strongly self-enforcing.*

Note that definition 2 is not equivalent to the statement that a sequence $J = \{J_1, J_2\}$ is stable if and only if it is stable in all $t$. However, it is straightforward that $J_2$ must be stable.

**Lemma 2** *In any stable sequence of coalitions $J = (J_1, J_2)$, $J_2$ is stable.*

Proof: We use a backwards induction argument. Suppose that $J_2$ is unstable with respect to $J'_2$. Then $J_t \rightarrow_M J'_t$ is strongly self-enforcing. Since $t = 2$ is the last period, it must be that $J \rightarrow_M J'$, where $J' = (J_1, J'_2)$ must be strongly self-enforcing for all $J_1$ in the larger game, too. Hence, $J$ cannot be stable. $\|$
From lemma 1 it becomes clear that our concept of stability is weaker than both $SNE$, and $CPNE$, since any deviation, which is subject to further deviation by any $J \in \mathbf{J}$ is not valid. Both $SNE$, and $CPNE$ allow for such deviations. [7]

# 4 Stable coalitions

In this section, we examine first the case of a single decision, i.e. the reduced two-stage game, in which a coalition is formed once. This not only allows for a comparison with the results of traditional theories of coalition formation in a single dimensional policy space but, by definition 2, is also a full description of the second stage in the sequential coalition formation game.

[7]Our definition also relates to the cooperative solution of the core of the underlying cooperative non-transferable utility (NTU) game (see e.g. Banerjee et al. [4]). Recall that a coalition $J$ is in the core of its underlying NTU game if and only if there is no non-empty coalition $J'$ such that for all $i \in J'$, $J' \succeq_i J'$ and $J' \succeq_j J$ for some $j \in J'$, which clearly relates to our definition 2i).

## 4.1 A single decision

A single decision is a decision on a single-dimensional policy space. Single-dimensionality of the policy space implies that different political parties can be (strictly) ordered on the real line or some interval on the real line. The traditional interpretation of this is that of an ideological left-right spectrum of political parties. Let $x_1 < x_2 < x_3$.

**Proposition 1** *In a single decision problem there is always a non-degenerate stable coalition* $J \in \mathbf{J}$.

Proof: First note that the grand coalition (123) is never dominated by the disconnected coalition (13), since (123) is always either strictly preferred to (13) by one of the two parties, or weakly preferred by both (if $S(13) = S(123)$). Hence, for at least one member of (13) there is no incentive to jointly deviate from (123).
Next, (12) and (23) are always stable with respect to (123), since it must be that $(12) \succ_1 (123)$ and $(23) \succ_3 (123)$. Hence, in order for not having a stable coalition, it must be that (12) and (23) are not stable with respect to each other. But since the median voter party is a member of both (12) and (23), this requires that $(12) \succ_2 (23)$ and $(23) \succ_2 (12)$, a contradiction. Thus, there is always at least one stable coalition. $\|$
Intuitively, this existence results is driven by the implicit single-peaked character of parties' preferences. Clearly, without such an assumption, stable coalitions do not need to exist (see e.g. Brams et al. [9]). However, due to the definition of stability, non-single-peaked preferences such as $(12) \succ_1 (13) \succ_1 (123)$, $(23) \succ_2 (123) \succ_2 (12)$, and $(13) \succ_3 (23) \succ_3 (123)$ do generate stable coalitions. Although for all $J$ there are self-enforcing deviations $J \rightarrow_M J'$ ($(12) \rightarrow_{(23)} (23)$, $(23) \rightarrow_{(13)} (13)$, $(13) \rightarrow_{(23)} (23)$, $(123) \rightarrow_{M \neq \{1,2\}} J, J \neq (23)$), none of these deviations are strongly self-enforcing, since self-enforcing deviations are cyclical. That means that, in such a case, all non-degenerate coalition structures are stable. Note that, however, these preferences are ruled out by the underlying bargaining solution of the game, as it is shown below. The next result, a corollary to proposition 1, shows that degenerate coalition structures can a priori be ruled out as candidate solutions for stable coalitions.

**Corollary 2** *Any degenerate coalition is unstable with respect to some non-degenerate coalition.*

Proof: Any degenerate coalition structure is induced by a strategy vector $\mathbf{s} = (s_1, s_2, s_3)$ with $s_1 \neq s_2 \neq s_3$. By lemma 1 any deviation $\rightarrow J$, such that $s_i = s_j$ for some $i, j$ is Pareto improving and feasible. By proposition 1 at least one of these deviations must be self-enforcing and induce a stable coalition. ||

Obviously, a strategy vector $\mathbf{s} = (s_1, s_2, s_3)$ with $s_1 \neq s_2 \neq s_3$ can constitute a Nash equilibrium, since this only requires that there is no profitable deviation by some $i \in \{1, 2, 3\}$. For example, in a one-shot game, the strategy vector $(s_1, s_2, s_3) = (1, 2, 3)$ is an (inefficient) Nash equilibrium.

**Definition 3** *A coalition $J \in \mathbf{J}$ is said to be disconnected if there is a $j \notin J$ such that $|x_j - x_i| < |x_k - x_i|$ for all $i, k \in J$.*

In other words, a coalition is disconnected if not all members of the coalition are adjacent to one another, i.e. it creates a 'hole' in an otherwise connected coalition. In the case of only three parties, this means that a (non-degenerate) coalition is disconnected if and only if the median party is not a member of that coalition.

**Corollary 3** *i) In a single decision there is no stable disconnected coalition. ii) The median party is always the member of a stable coalition.*

Proof: It suffices to show that (13) is always unstable with respect to either (12) or (23). 2 lexicographically prefers both (12) and (23) to (13). Now suppose that $B(13) < B(12)$, that is, $(13) \succ_1 (12)$. Then it must be that $(23) \succ_3 (13)$ since $B(12) < x_2 < B(23)$. Conversely, by the same argument, if $(13) \succ_3 (23)$ then $(12) \succ_1 (13)$, which completes the proof of the first statement. The second statement then follows trivially. ||

Corollary 3 re-establishes the well known result of political coalition formation in a single dimensional policy space that all stable coalitions must be connected. Consequently, the median party must always be the member of a stable coalition. However, in contrast to previous results, it is not only minimal winning coalitions that are stable. Generically, there is either a unique stable coalition, consisting of two members, or there are two stable coalitions, one of which has to be the grand coalition, i.e., if there is no unique stable coalition, then the grand coalition must be stable.

**Lemma 3** *In a single decision, if the grand coalition is unstable with respect to some coalition $(ij)$ then $(ij)$ is the unique stable coalition.*

Proof: First, let (123) be dominated by (12), i.e., $B(123) > x_2$ and $|B(123) - x_2| > |B(12) - x_2|$. Then (12) must be stable since $B(23) > B(123)$ and thus $(12) \succ_2 (23)$. By corollary 3 (12) must also be stable with respect to the disconnected coalition (13). Hence, (12) is the unique stable coalition. Conversely, if (123) is dominated by (23) then by the same argument as above, (23) is the unique stable coalition. ||

We now establish necessary and sufficient conditions for the existence of a unique stable coalition.

**Proposition 4** *Let* $x_1 < x_2 < x_3$*, i.e., 2 is the median party. Then there is a unique stable coalition if and only if*

$$\frac{p_i(1-p_j)+p_i}{p_j(1-p_j)} < \frac{d(j2)}{d(i2)}, \quad \{i,j\} = \{1,3\} \tag{1}$$

Proof: From lemma 3 there is a unique stable coalition if and only if either $(12) \succ_2 (123)$ or $(23) \succ_2 (123)$, which implies that $|B(123)-x_2| > |B(12)-x_2|$. We distinguish two cases with two subcases each.
Case 1: $B(12) \succ_2 B(23) \Rightarrow (12)$ is stable by proposition 1.
If (12) is the unique stable coalition, then $B(123) > x_2$. Hence, $p_1 d(12)+p_3 d(23) > 0$, and we require

$$\frac{p_1}{p_1+p_2} d(12) < -p_1 d(12) + p_3 d(23)$$

which, since $(p_1+p_2) = (1-p_3)$ gives

$$\frac{p_1(1-p_3)+p_1}{p_3(1-p_3)} < \frac{d(23)}{d(12)}$$

i.e., $i=1$, and $j=3$ in equation 1.
Case 2: $B(23) \succ_2 B(12) \Rightarrow (23)$ is stable by proposition 1.
If (23) is the unique stable coalition, then $B(123) < x_2$. Hence, $-p_1 d(12)+p_3 d(23) < 0$, and we require

$$\frac{p_3}{p_2+p_3} d(23) < p_1 d(12) - p_3 d(23)$$

which, since $(p_2+p_3) = (1-p_1)$ gives

$$\frac{p_3(1-p_1)+p_3}{p_1(1-p_1)} < \frac{d(12)}{d(23)}$$

i.e., $i=3$, and $j=1$ in equation 1, which completes the proof. $\|$
The necessary and sufficient condition 1 dictates that a stable coalition is unique if either two parties are ideologically relatively close to each other, of if the median party is very strong compared to the party it wishes to form a coalition with.
In order to illustrate the connection between Nash equilibria, coalition proofness and stability, consider the following example of a single-decision model.

**Example 1** *Let* $x_1 < x_2 < x_3$*, and let* $(123) \succ_2 (23) \succ_2 (12)$*, i.e., both the grand coalition* $(123)$ *and* $(23)$ *are stable. The strategy set is given by* $S^i = \{i, ij, ijk\},\ i,j,k = 1,2,3$*. Denote by* $(s_1, s_2, s_3)$ *a strategy vector,* $NE$ *the set of all Nash equilibria, and* *by $STABLE$ the set of strategy profiles which induce stable coalitions. Then*

$$\begin{aligned} NE &= \{(1,2,3),(1,2,123),(1,123,3),(123,2,3), \\ &\quad (s_1 \in S^1, 23, 23), (123,123,123), (12,12,123), \\ &\quad (12,12,3),(12,12,13),(13,2,13),(13,123,13)\} \\ STABLE &= \{(s_1 \in S^1, 23, 23), (123,123,123)\} \end{aligned}$$

*The equilibrium outcome of the four $(s_1 \in S^1, 23, 23)$ is $(23)$. Note that any strategy $s_1 \in S^1$ is maximal, since party $1$ is indifferent between all policy outcomes if it is not member of the coalition. Note that, while all non-degenerate coalitions are efficient, not all of them are Nash equilibria. $\mathbf{s} = (13, 23, 13)$ is clearly Pareto efficient, but since there is a unilateral beneficial deviation for party $3$ ($s'_3 = 23$), it is not a Nash equilibrium. On the other hand, there are efficient Nash equilibria, which do not induce stable coalitions. For example, $s = (13, 2, 13)$ is efficient, but $(s'_2, s'_3) = (23, 23)$ is clearly a profitable joint deviation.*

As explained above, in a single decision case the solution concepts $SNE$, $CPNE$, and stability induce the same equilibrium outcomes.

**Remark 1** *In a single decision problem $SNE \Leftrightarrow STABLE$ and $CPNE \Leftrightarrow STABLE$.*

Proof: Since $SNE \Rightarrow CPNE$ it suffices to show that $CNPE \Rightarrow STABLE \Rightarrow SNE$. First, since strong self-enforceability strictly includes all deviations of proper sub-coalitions of deviating coalitions, $CPNE \Rightarrow STABLE$.
For the second part, recall that $\mathbf{s}$ is a strong Nash equilibrium if and only if for all $J \in \mathbf{J}$ and for all $\mathbf{s}'_J = \{s_j\}_{j \in J}$ there is an $i \in J$ such that $\Psi(\mathbf{s}) \succeq_i \Psi(\mathbf{s}'_J, \mathbf{s}_{-J})$. Now suppose that $\Psi(\mathbf{s}) = J$ is stable, i.e. either $J = (123)$ or $J = (2i)$ with $i \in \{1, 3\}$. For both case we have to consider every conceivable coalitional deviation. Consider first $J = (123)$. Let $\mathbf{s} = \Psi^{-1}(J)$. Clearly, for $J' \in \{(123), (12), (23), \}$, we have that $\Psi(\mathbf{s}) \succeq_2 \Psi(\mathbf{s}'_J, \mathbf{s}_{-J})$. For $J' = (13)$, either 1 or 3 strictly prefer $\Psi(\mathbf{s})$ to $\Psi(\mathbf{s}'_{13}, \mathbf{s}_2)$, and finally $\Psi(\mathbf{s}) \succ_i \Psi(\mathbf{s}'_i, \mathbf{s}_{\{-i\}})$ for all $J' = (i)$. Hence, $\Psi(\mathbf{s})$ is a strong Nash equilibrium. Now let $J = (i2)$. For $J' = (123)$, $\Psi(\mathbf{s}) \succ_i \Psi(\mathbf{s}'_{123})$, for $J' = (2j)$, $j \neq i$ both 2 and $i$ prefer $\Psi(\mathbf{s})$, for $J' = (13)$ it holds for 2, and finally $j \neq i$ weakly prefers $\Psi(\mathbf{s})$ to $\Psi(\mathbf{s}'_j, \mathbf{s}_{\{-j\}})$, which completes the proof. $\|$
The single decision problem is the building block for our further analysis, which consists of two quite different coalition formation games over two decisions.

## 4.2 Two decisions

While, for obvious reasons, this assumption of a single-dimensional policy space is hardly controversial in two-party systems, it is by far less clear whether a single dimensional policy space is appropriate in contemporary multi-party systems. Frequently, political parties take different positions relative to each other on different political issues.[8] A straightforward solution to this problem is the introduction of a multi-dimensional policy space, in which parties are situated, so that the favorite policy of a party is described by a vector rather than a number. Clearly, a multi-dimensional policy space deprives the term of connected coalitions, at least partially, of its meaning. Nevertheless, ideological proximity can be described as the distance between the location of different parties. Kirchsteiger and Puppe [15] generalize the classical models in this direction. Their main results only partially confirm the older literature, since stable coalitions in the presence of many parties turn out to exist only under restrictive assumptions. Still, if stable coalitions exist (which is more likely under the assumption of office- and policy-seeking parties), they are generally minimum winning coalitions. In the fist part of this section, we explore the multi-dimensional case in our set-up.

In the second part of the section we challenge a second fundamental assumption, which is less explicit than the dimension of the policy space. Political coalition formation games are usually seen as one-shot games, which implies that either party positions do not change over the legislative session (that is, over a sequence of decisions), or that only a single coalition can be formed over several policy issues.[9] While the introduction of a multi-dimensional policy space takes into account the multiplicity of decisions, given that each dimension can be seen as one decision problem that requires a solution, it clearly requires that all decisions are made simultaneously and within the same coalition.

Arguably, simultaneous decision-making and the dimension of the policy space are closely related. Indeed, if one interprets each dimension in a multi-dimensional policy space as an autonomous decision problem, they are of "dual" character, since, as we show below, repeated coalition formation with fixed positions on a single-policy space cannot change the coalitional structure of the single decision, while, on the other

[8]Kirchsteiger and Puppe [15] give as an example the green parties in Europe, but even the more traditional parties do not fit the assumption of single-dimensionality. The vertical integration of the European Union; in respect to which left and right wing parties share similar, though differently-motivated reservations provides an appropriate example.

[9]Alternatively, coalition formation games can be modelled in a sequential game, if e.g. a bargaining process is involved. Obviously, such games are not one-shot. However, often only a single coalition is to be formed, see e.g Bloch [7].

hand, any sequential coalition formation without positions being fixed over time, is to be seen as multi-dimensional policy space.

### 4.2.1 Single coalitions over two decisions

If a coalition can only be formed once for all decisions, then the coalition formation problem is equivalent to a single decision in a two-dimensional policy space. Graphically, the location of ideal policies of the parties form a triangle in $R_+^2$, within which we find the bargaining solution. While all solutions to minimal winning coalitions lie on the vertices of that triangle, the solution to the grand coalition (123) lies in its interior. Essentially, if only one coalition can be formed, the analysis of a single decision carries over to the multiple decision case. The distance between two parties is now given by $d(ij) = \sqrt{d_1^2(ij) + d_2^2(ij)}$.
Consequently, we summarize the main result in a single proposition.

**Proposition 5** *Generically, in a coalition formation problem with two-dimensional policy space, there are, at most, two stable coalitions. If there is a unique stable coalition, it is a minimal winning coalition. If there are two stable coalitions, the grand coalition* (123) *is always stable.*

Proof: First, we show by contradiction that there is always a stable minimal winning coalition. Suppose that $(12) \succ_1 (13), (23) \succ_2 (12)$, and $(13) \succ_3 (23)$, i.e. there is no stable minimal winning coalition. Then

$$(12) \succ_1 (13) \quad \Rightarrow \quad p_2(1-p_2)d(12) < p_3(1-p_3)d(13) \tag{2}$$

$$(23) \succ_2 (12) \quad \Rightarrow \quad p_3(1-p_3)d(23) < p_1(1-p_1)d(12) \tag{3}$$

$$(13) \succ_3 (23) \quad \Rightarrow \quad p_1(1-p_1)d(13) < p_2(1-p_2)d(23) \tag{4}$$

Solving for $d(13)$ in (2) and for $d(23)$ in (3) and substituting in (4) gives

$$p_1(1-p_1)\frac{p_2(1-p_2)}{p_3(1-p_3)}d(12) < p_2(1-p_2)\frac{p_1(1-p_1)}{p_3(1-p_3)}d(12) \tag{5}$$

which reduces to $d(12) < d(12)$, a contradiction. Clearly, the stable minimal winning coalition is unique, since for all $i$, we have either $(ij) \succ_i (ik)$ or $(ik) \succ_i (ij)$, $i \neq j \neq k$
Next we show that the grand coalition can indeed be stable. Let (12) be stable, i.e.,

$$(12) \succ_1 (13) \quad \Rightarrow \quad \frac{p_2(1-p_2)}{p_3(1-p_3)} < \frac{d(13)}{d(12)} \tag{6}$$

$$(12) \succ_2 (23) \quad \Rightarrow \quad \frac{p_1(1-p_1)}{p_3(1-p_3)} < \frac{d(23)}{d(12)} \tag{7}$$

For (123) to be stable, either $(123) \succ_1 (12)$, or $(123) \succ_2 (12)$. W.l.o.g. suppose $(123) \succ_1 (12)$, which implies that

$$(p_2 + p_3)d(1B(23)) < \frac{p_2}{p_1 + p_2} d(12) \tag{8}$$

where $d(1B(23))$ is the distance between party 1 and the bargaining solution $B(23)$ which by a geometrical argument and with the application of some algebra is given by

$$d(1B(23)) = \sqrt{d^2(12) + \left[\frac{p_3}{p_2 + p_3} d(23)\right]^2 - \left(\frac{p_3}{p_2 + p_3}\right)(d(12)^2 + d^2(23) - d^2(13))}$$

such that condition (8) can be written as

$$\left[p_2^2 \left(1 - \frac{1}{(1 - p_3)^2}\right) + p_2 p_3\right] d^2(12) + p_3(p_2 + p_3)d^2(13) < p_2 p_3 d^2(23) \tag{9}$$

Since we are free to choose any value for $d(23)$ in (7) and (9), we observe that for large enough values of $d(23)$, inequalities (6),(7), and (9) can indeed hold simultaneously, we conclude that (123) can be stable. Moreover, since it is shown above that there is always exactly one stable minimal winning coalition, it must be that (123) is never the unique stable coalition, which completes the proof. ||

This result contrast with Kirchsteiger and Puppe [15], who find that the grand coalition is never stable, because their solution depends on the sum of the distances (irrespective of the direction) between potential coalition members, so that parties always prefer a smaller to a larger coalition, maybe thereby reflecting bargaining costs, which increase as the the number of parties involved increase. However, if bargaining costs are low, it is intuitively correct that a party in the political center might prefer a compromise between all parties to any other compromise.

While in the single decision case there are disconnected coalitions, which are never stable, here all non-degenerate coalitions can be stable for some parameter values. However, a coalition $J$ with $d(J) > d(J')$ for some $J'$ can only be stable for some values of $p_1, p_2, p_3$ if $J'$ is stable for some $p_1', p_2', p_3'$. Conversely, if $J$ is stable for some parameter values, then this must be the case for $J'$, too.

**Proposition 6** *Let $d(J) > d(J')$. Then $J$ can be stable for some $p_1, p_2, p_3$ only if $J'$ is stable for some $p_1', p_2', p_3'$. Conversely, if $J$ is stable for some $p_1, p_2, p_3$, then $J'$ is always stable for some $p_1', p_2', p_3'$*

Proof: We start with the second part. Let $d(13) > d(23) > d(12)$, and let (13) be stable for $p_1, p_2, p_3$, i.e

$$\frac{p_3(1 - p_3)}{p_2(1 - p_2)} < \frac{d(12)}{d(13)}, \text{ and } \frac{p_1(1 - p_1)}{p_2(1 - p_2)} < \frac{d(23)}{d(13)}$$

For (23) to be stable, we need

$$\frac{p_3(1-p_3)}{p_1(1-p_1)} < \frac{d(12)}{d(23)}, \text{ and } \frac{p_2(1-p_2)}{p_1(1-p_1)} < \frac{d(13)}{d(23)}$$

Since by assumption $\frac{d(12)}{d(13)} < \frac{d(12)}{d(23)}$, and $\frac{d(23)}{d(13)} < \frac{d(13)}{d(23)}$, coalition (23) must always be stable for $p_1' = p_2, p_2' = p_1, p_3' = p_3$.
The first part of the statement now follows straightforwardly. $\|$
However, the following proposition shows that the results of the single decision carries over to the two decision case if the median party is the same in both decisions, i.e. this party will always be a member of all stable coalitions.

**Proposition 7** *If a party $i$ is the median party in both dimensions, then a coalition $J$ with $i \notin J$ is never stable.*

Proof: We use a contradiction argument in order to show this result. Let 2 be the median party in both dimensions. This implies that $d^2(12) + d^2(23) < d^2(13)$. Then, substituting for $d(13)$ in the usual stability conditions, for (13) to be stable the inequalities

$$\begin{aligned}
\left[\frac{p_3(1-p_3)}{p_2(1-p_2)}\right]^2 [d^2(12) + d^2(23)] &< d^2(12) \\
\left[\frac{p_1(1-p_1)}{p_2(1-p_2)}\right]^2 [d^2(12) + d^2(23)] &< d^2(23)
\end{aligned}$$

must hold, which after some algebra reduces to

$$\frac{p_1(1-p_1)}{p_2(1-p_2) - p_1(1-p_1)} < \frac{d(12)}{d(23)} < \frac{p_2(1-p_2) - p_3(1-p_3)}{p_3(1-p_3)}$$

By substituting for $p_3$ we find a condition which must hold for some $p_1, p_2 \leq \frac{1}{2}$.

$$\frac{p_1(1-p_1)}{p_2(1-p_2) - p_1(1-p_1)} < \frac{p_2(1-p_2) - (1-p_1-p_2)(p_1+p_2)}{(1-p_1-p_2)(p_1+p_2)} \tag{10}$$

Recall that, since $d(23) < d(13)$, for (13) to be stable we have to have $p_1 < p_2$, which implies that the left hand side of (10) is positive, and the inequality can be rewritten as

$$p_1(1-p_1) + (1-p_1-p_2)(p_1+p_2) - p_2(1-p_2) < 0$$

which further reduces to

$$2p_1(1-p_2-p_1) = 2p_1p_3 < 0$$

which, since $p_1, p_3 > 0$, is a contradiction. Hence, coalition (13) is never stable. $\|$
The result is quite intuitive, since if one party is the median party in all dimension, we do expect it to be a member of the stable coalition.[10] Moreover, if the two-dimensional space is split into two single-dimensional policy spaces, a coalition (13) means that a disconnected coalition can be stable. Indeed, in the following section, we show that this cannot be the case in a sequential coalition formation game, either.

### 4.2.2 Sequential decision-making

If coalition formation is assumed to be a sequential process, then parties are allowed to talk not only before the first decision is to be made, but also between each decision. This, obviously, enlarges the strategy space of political parties. Nash equilibria can be time-inconsistent. However, the associated strategy vector of any stable sequence of coalitions, as in definition 2, is time consistent.
In the following paragraphs, we describe some central results regarding the set of stable sequences of coalitions. Not surprisingly, under different parameter restrictions, the set of stable sequences of coalitions exhibits different characteristics. However, stable sequences of coalitions will always exist, although the members of the coalitions can change over time.[11]
W.l.o.g. we will always assume that $x_{11} < x_{21} < x_{31}$. Moreover, by symmetry, in the second decision we only have to distinguish two cases. In the first case, the median party is the same in both decisions, that is, $x_{12} < x_{22} < x_{32}$, which is equivalent to $x_{32} < x_{22} < x_{12}$. In the second case, the median party in the second decision is different to the median party in the first decision. Again, by symmetry, w.l.o.g. we assume that $x_{22} < x_{12} < x_{32}$. All other possible cases, $x_{12} < x_{32} < x_{22}$, $x_{32} < x_{12} < x_{22}$, and $x_{22} < x_{32} < x_{12}$, are completely symmetric to this one.

**Proposition 8** *If there is a unique stable coalition in the second decision, then any stable sequence of coalitions $J = (J_1, J_2)$ is stable in each single decision.*

Proof: Clearly, in the second period, parties will form the unique stable coalition. By a backwards induction argument, $J_1$ has to be stable, too. Since parties know that $J_2$ is the unique stable coalition in the second decision, no member of any stable coalition $J_1$ can have an incentive to deviate in order to induce another outcome than

[10] Note that the result holds for a wider class of decision problems, because while the assumption of a fixed median party implies $d^2(12) + d^2(23) < d^2(13)$, the converse statement does not hold.

[11] Arguably, in this case, the term "stable" is misleading, since if the members of a coalition change within a legislative session, one can hardly speak of the "stability" of that coalition. Nonetheless, this is fully compatible with the solution concept.

$J_2$ in the second decision. Therefore, parties will form a stable coalition in the first decision, too. $\|$

Proposition 8 dictates that, if there is a unique stable coalition in the second decision, then there is no way to select among any coalition in the first period. This means that any sequence of stable coalitions can turn out to be stable. In particular, and in contrast to simultaneous decision-making, coalitions can change over time, i.e., different coalitions are formed in each decision. The following corollary is an immediate consequence of proposition 8.

**Corollary 9** *If there is a unique stable coalition in the second decision, then the median party is always a member of both $J_1$ and $J_2$.*

In particular, corollary 9 also implies that if there is an $i \in \{1, 2, 3\}$ which is the median party in both decisions, then this party will always be a member of both coalitions. Clearly, in contrast to the single coalition formation case, the sequence of coalitions may consist of different coalitions in the two decisions.

Proposition 8 only partially carries over to the situation in which there is no unique coalition in the second period. While in the former case, a sequence of coalitions is stable if and only if it is stable in every period, in the latter case, sequences of stable coalitions are not the only stable sequences. However, the following result shows that still every sequence of stable coalitions can be stable, i.e. stability in every period is a sufficient, but not necessary, condition.

**Proposition 10** *Any sequence of period-stable coalitions $J = (J_1, J_2)$ is stable.*

Proof: Let $(s) = \Psi^{-1}(J)$. By lemma 2 $J_2$ is stable. Now let $(s)$ be such that
i) $s_{i1} = J_1$ for all $i \in J_1$, and $s_{j1} \in \mathbf{J}_j$ for $j \notin J_1$.
ii) For all $J_1' \in \mathbf{J}$, $s_{i2}(J_1') = J_2(J_1')$ for all $i \in J_2$ and, and $s_{j2}(J_1') \in \mathbf{J}_j$ for $j \notin J_2$.
Then there is no $M \subset \{1, 2, 3\}$ such that $J \rightarrow_M J'$ is strongly self-enforcing, since for all $M$ at least one $i \in M$ strictly prefers $J$ to any $J'$. Hence $J$ is a stable sequence of coalitions. $\|$

As noted above, sequences of stable coalitions are not the only stable sequences, if 123 is stable in the second decision, i.e. there is no unique stable coalition. It turns out that in a stable sequence of coalitions, $J_1$ not only does not have to be stable itself, but is can also be disconnected.

**Proposition 11** *Let $x_{22} < x_{12} < x_{32}$. If the grand coalition* (123) *is stable in the second decision, then there are stable sequences of coalitions $J = (J_1, J_2)$ where $J_1$ is* *not stable. In particular, $J_1$ can be disconnected if and only if*

$$\begin{aligned}(13_2) &\succ_1 (12_2)\\(123_2) &\succ_1 (13_2)\\(23_1) &\succ_2 (12_1)\\(13_1, 13_2) &\succ_3 (23_1, 123_2)\end{aligned}$$

*and $J_1$ can be connected, but unstable if and only if*

$$\begin{aligned}(12_2) &\succ_1 (13_2)\\(123_2) &\succ_1 (12_2)\\(23_1) &\succ_2 (12_1)\\(12_1, 12_2) &\succ_3 (23_1, 123_2)\end{aligned}$$

Proof: Let $x_{22} < x_{12} < x_{32}$, i.e., there is an $i \in \{2, 3\}$ such that $J_2 = (1i_2)$ is stable, and let $(123_2) \succ_1 (1i_2)$. Moreover, let $(23_1)$ be stable in decision one, i.e. $(23_1) \succ_2 (12_1)$. We construct a strategy vector $\mathbf{s}$ with $\Psi(\mathbf{s}) = ((1i), (1i))$.
i) $s_1 = (s_{11}, s_{12})$: $s_{11} = (1i)$, $s_{12} = (1i(1i_1), 123(J_1')))$, where $J_1' \neq (1i)$, i.e. party 1 plays $(1i)$ in the first decision, and in decision 2 it plays $(1i)$ if it has been formed in the first decision. If $(1i_1)$ is not formed in $t = 1$ (i.e. $s_i \neq (1i)$, it plays $(123)$ in $t = 2$.
ii) $s_i = (s_{i1}, s_{i2})$: $s_{i1} = (1i_1)$, $s_{i2} = (1i(J_1)))$ for all $J \in \mathbf{J}$.
iii) $s_{j \notin (1i)} = (s_{j1}, s_{j2})$: $s_{i1} = (J_1' \in \mathbf{J_j})$, $s_{i2} = (J_2(J_1),\ J_2 \in \mathbf{J}_j)$ for all $J \in \mathbf{J}$.
Now suppose that $(1i_1, 1i_2) \succ_i (23_1, 123_2)$ for some $i \neq 1$. Then there is no $M \in \mathbf{J}$ such that $J \rightarrow_M J'$ is strongly self-enforcing. Clearly there is no $M$ with $i \in M$, which has an incentive to deviate in a self-enforcing way. For $M,\ 1 \in M$, notice that $J \rightarrow_{(1j), j \neq i} (1j_1, 123_2)$ can be self-enforcing, if $(1j_1, 123_2) \succ_1 (1i_1, 1i_2)$. However, it can never be strongly self-enforcing, since, if $(23_1)$ is stable, then $(1j_1, 123_2) \rightarrow_{(23)} (23_1, 123_2)$ must be self-enforcing. Hence $J = ((1i_1), (1i_2))$ is stable, and $J_1$ is disconnected if $i = 3$, and connected, but unstable in $t = 1$, if $i = 2$. ‖
Proposition 11 says that even if (23) is stable in decision 1, there are stable sequences of coalitions, in which other minimal winning coalitions are formed in the first decision. In particular, it is shown above that disconnected coalitions can be sustained in the sequential coalition formation game. Clearly, this holds a fortiori for connected sequences.
The necessary and sufficient conditions for the existence of such a partially discon-

nected sequence are given by

$$\frac{p_3}{1-p_2}d_2(13) \quad < \quad \frac{p_2}{1-p_3}d_2(12) \tag{11}$$

$$|p_3 d_2(13) - p_2 d_2(12)| \quad < \quad \frac{p_3}{1-p_2}d_2(13) \tag{12}$$

$$\frac{p_3}{(1-p_1)}d_1(23) \quad < \quad \frac{p_1}{1-p_3}d_1(12) \tag{13}$$

$$\frac{p_1}{1-p_2}(d_1(13)+d_2(13)) \quad < \quad \frac{p_2}{1-p_1}d_1(23) + p_1 d_2(13) + p_2 d_2(23) \tag{14}$$

From inequality (12) we distinguish two cases, which induce mutually exclusive conditions (11)-(12).

Case 1: $p_3 d_2(13) - p_2 d_2(12) < 0$:

$$\frac{d_2(12)}{d_2(13)} \in \left( \max\left\{ \frac{p_3}{p_2}, \frac{p_3(1-p_3)}{p_2(1-p_2)} \right\}, \frac{p_3 + p_3(1-p_2)}{p_2(1-p_2)} \right) \tag{12.1}$$

Case 2: $p_3 d_2(13) - p_2 d_2(12) > 0$:

$$\frac{d_2(12)}{d_2(13)} \in \left( \frac{p_3(1-p_3)}{p_2(1-p_2)}, \frac{p_3}{p_2} \right) \quad \Rightarrow \quad p_3 > p_2 \tag{12.2}$$

Clearly, inequalities (12.1) and (13) can hold simultaneously, as can (12.2) and (13). Then there must exist values of $d_t(ij)$, $t = 1, 2,\ i, j \in \{1, 2, 3\}$, such that inequality (14) holds, since for all $d_2(23) > 0$ there are $d_2(12)$ and $d_2(13)$, such that (12.1), and (12.2) respectively, hold. To see this, rearrange (14), such that both $d_2(13)$ and $d_2(23)$ only appear the right hand side of the inequality, and substitute $d_2(23)$ by $d_2(13) + d_2(12)$. Taking the total differential gives

$$(p_1 + p_2)\Delta d_2(12) + \left( p_2 - \frac{p_1}{1-p_2} \right) \Delta d_2(23) \tag{15}$$

which is always positive, since if $(p_2 - \frac{p_1}{1-p_2})$ is positive, it is clearly positive, and if $(p_2 - \frac{p_1}{1-p_2})$ is negative, the expression is positive if $(p_1 + p_2) > (\frac{p_1}{1-p_2}) - p_2$, which indeed holds for all for all $p_1, p_2 \in (0, \frac{1}{2})$ Hence $d_2(23)$ can be chosen freely, such that inequality (14) holds.

Note that such equilibria are more likely the larger the maximal difference between two parties' ideal policy in the second decision, compared to the maximal distance in the first decision, i.e $\arg\max_{ij}\{d_{(ij_2)}\}/\arg\max_{ij}\{d_{(ij_1)}\}$. Intuitively, this means that a disconnected coalition is more likely, if it first decides upon a rather uncontroversial policy issue, while the second decision, in which a connected coalition is formed, is more controversial. Arguably, this is not surprising, since if the second decision is rather controversial, a party is more willing to give up some opportunity in the first decision in order to benefit in the second one.

# 5 Comparing the results

We have taken the single decision coalition formation as a benchmark case. Most of the results of this case carry over to the single coalition formation case over two decisions, with the main difference being that in the latter case all non-degenerate coalitions can be stable under some allocation of bargaining powers. In particular, a coalition can be stable for some allocation of bargaining powers if and only if all ideologically closer coalition are stable for some other allocation of bargaining power. This is not true for the sequential game. Let $x_{11} < x_{21} < x_{31}$, and $x_{22} < x_{12} < x_{32}$. It is shown above that $J = (13_1, 13_2)$ can be stable if $(13)_2$ is stable. In particular, it can also be stable if $d_1(13) + d_2(13) < d_1(12) + d_2(12)$. However, if $(13)_2$ is stable, then $(12_2)$ cannot be stable and so cannot $J' = (12_1, 12_2)$.

If a party is the median one in all dimensions, then it will always be in the coalition, which holds true also for the sequential game. Comparing the two regimes in the case of a fixed median party shows that the median party (if it was able to choose) would always prefer a sequential coalition formation to a single one, since the possibility of repeated coalition formation improves at least weakly the median party's situation. Clearly, there is no such relation between the two regimes for any other party, because if one non-median party is better off in the sequential case, it must be that the other non-median party is worse off in that case.

There is a more general result in a similar spirit, which immediately follows from the fact that all non-degenerate coalitions and hence all sequences of non-degenerate coalitions are efficient with respect to the entire feasible payoff space.

**Proposition 12** *Stable coalitions and stable sequences of coalitions cannot be Pareto ranked.*

We also showed that in the sequential case coalition structures, which are unstable in the single decision case, can be supported as stable sequences of coalitions. In particular, disconnected coalitions can be induced by punishment strategies. Note that such sequences can be, but need not be perfectly coalition proof nor a fortiori perfectly strong Nash equilibria. Let $x_{11} < x_{21} < x_{31}$, and $x_{22} < x_{12} < x_{32}$, and let $J = (13_1, 13_2)$ be stable. Then $J$ is perfectly coalition proof if and only if $(13_1) \succ_1 (12_1)$, because otherwise $(13_1) \rightarrow_{\{1,2\}} (12_1)$ is a profitable joint deviation. However, sequences of stable coalitions, which are shown to be always in the set of stable sequences of coalitions, can always be supported by perfect coalition proof and perfectly strong Nash equilibrium. For the single coalition formation game the equivalence of the three solution concept is shown above.

While the single coalition formation game does not allow for re-negotiation, the sequential formation does. Hence, in the former setting, timing does not play any role, and parties form a coalition which is stable over the entire policy space. In the latter case, timing obviously plays an important role, which does not always allow for some overall beneficial coalition formation. In particular, stable sequences, which contain otherwise unstable coalitions, are never robust with respect to changes in the order of decisions.

**Proposition 13** *Stable sequences which contain otherwise unstable coalitions, are never robust with respect to changes in the order of decisions.*

To see this, let $x_{11} < x_{21} < x_{31}$, and $x_{22} < x_{12} < x_{32}$, and let $J = (13_1, 13_2)$ be stable. Reversing the order, it can never be that $J = (13_1, 13_2)$ remains stable, since $(13_2)$ can never be stable in the reversed order. On the other hand, sequences of stable coalitions never depend on the order of decisions, since the same set of coalitions can be supported in reverse order.

# 6 Conclusion

We introduced a general concept of stable sequences of coalitions with clear similarities to the previously introduced non-cooperative solution concept of (re)negotiation-proof Nash equilibria. This solution concept is then applied to the three different games of political coalition formation, where the one-shot coalition formation over a single decision is the constituent of both the single coalition formation over two decisions and the sequential coalition formation game. It turns out that our model does not only allow for the same coalition outcomes as previous theories of political coalition formation, but the sequential game also allows for richer coalition structures, in particular disconnected coalitions in the non-final stage and, more generally, otherwise unstable coalition structures. The larger set of possible stable coalition structures also induces a multiplicity of equilibria within a wide range of parameter restrictions and hence does not always allow for clear predictions of the outcome of the coalition formation process, since there is no natural equilibrium selection criterion.
There are two natural extension to our analysis, which both can be seen as agenda setting issues. Firstly, for a given allocation of bargaining powers, there are winners and losers in each regime. Consequently, the choice of the regime can be subject to a bargaining problem itself. Secondly, in the previous section it is shown that the order of the decisions is important for the type of equilibrium that emerges. This exogenous

timing of the decision can be seen as a given agenda of policy issues. Hence, another natural extension of our analysis is the endogeneization of the agenda setting process in the sequential coalition formation game.[12]

# References


[1] A. Alesina, and H. Rosenthal, Partisan politics, divided governmant, and the economy, Cambridge, 1995.

[2] R. Aumann, Acceptable points in cooperative general n-person games, in "Contributions to the theory of games IV", Princeton, 1959.

[3] R. Axelrod, Conflict of interest, Chicago, 1970.

[4] S. Banerjee, H. Konishi, and T. Sönmez, Core in a simple coalition formation game, *Social Choice and Welfare* 18 (2001), 135-153.

[5] D. Baron, and J. Ferejohn, Bargaining and agenda formation in legislatures, *American Economic Review* 77 (1987), 303-309.

[6] B. D. Bernheim, B. Peleg, and M. D. Whinston, Coalition-proof Nash equilibria I. Concepts, *Journal of Economic Theory* 42 (1987), 1-12.

[7] F. Bloch, Sequential formation of coalitions in games with externalities and fixed payoff division, *Games and Economic Behavior* 14 (1996), 90-123.

[8] F. Bloch, and S. Rottier, Agenda control in coalition formation, *Social Choice and Welfare* 19 (2000), 769-788.

[9] S. J. Brams, M. A. Jones, and D. M. Kilgour, Single-Peakedness and disconnected coalitions, *Journal of Theoretical Politics* 14 (2002), 359-383.

[10] J. Bulow, and P. Klemperer, The generalized war of attrition, CEPR discussion paper 1561, 1997.

[11] F. Carmignani, The Representation of the Political System in the Political-Economy Literature, *International Review of Economics and Business* 47 (2000), 353-390.

[12] F. Carmignani, Cabinet formation in coaltion systems, Glasgow University Discussion Paper 2004 (2000).

[13] A. De Swan, Coalition theories and cabinet formation, Amsterdam, 1973.

[14] L. C. Dodd, Coalitions in parliamentary government, Princeton, 1976.

[15] G. Kirchsteiger, and C. Puppe, On the formation of political coalitions, *Journal of Institutional and Theoretical Economics*, 153 (1997), 293-319.


[12]Baron and Ferejohn [5] provide a model which analyzes agenda setting in a bargaining process. Bloch and Rottier [8] explore, in a different set up, under which circumstances certain equilibria can be imposed by strategic agenda control.

[16] J. von Neumann, and O. Morgenstern, The theory of games and economic behaviour, Princeton, 1953.

[17] W. Riker, The theory of political coalitions, New Haven, 1962.

[18] L. Xue, Negotiation-proof Nash equilibrium, *International Journal of Game Theory* 29 (2000), 339-357.