Endogenous Parties in an Assembly. The Formation of Two Polarized Voting Blocs. *

Jon X Eguia†
New York University

September 12, 2007

Abstract

In this paper I show how members of an assembly form voting blocs strategically to coordinate their votes and affect the policy outcome chosen by the assembly. In a repeated voting game, permanent voting blocs form in equilibrium. These permanent voting blocs act as endogenous political parties that exercise party discipline. In a stylized assembly I prove that the equilibrium parties must be two small polarized voting blocs. In an empirical application of the model to the US Supreme Court, I again predict that the equilibrium outcome of strategic coalition formation in the Court would lead to the formation of two voting blocs, one at each side of the ideological divide.

JEL: D71, D72.

*I thank Francis Bloch, Anna Bogomolnaia, Matias Iaryczower, Matt Jackson, Andrea Mattozzi and Tom Palfrey for their comments and suggestions. I am grateful to Andrew Martin and Keith Poole for their generous help with the US Supreme Court data.

†Email: eguia@nyu.edu. Mail: 19 West 4th Street, 2nd Floor, Dept. of Politics NYU. New York, NY 10012.

1 Introduction

Democratic deliberative bodies, such as committees, councils, or legislative assemblies across the world choose policies by means of voting. Members of an assembly can affect the policy outcome chosen by the assembly by coordinating their voting behavior and forming a voting bloc. A voting bloc is a coalition with an internal rule that aggregates the preferences of its members into a single position that the whole coalition then votes for, acting as a single unit in the assembly. From factions at faculty meetings in an academic department, to alliances of countries in international relations or political parties in legislative bodies, successful voting blocs influence policy outcomes to the advantage of their members. In national politics, legislators face incentives to coalesce into strong political parties in which every member votes according to the party line. Exercising party discipline to act as a voting bloc, strong parties are more likely to attain the policy outcomes preferred by a majority of party members.

However, agents are not identical and the benefits of forming a voting bloc are not equally shared by all. Some members of a voting bloc may prefer to leave the bloc, making it unstable. Who benefits when agents with diverse preferences form a voting bloc? What voting blocs do we expect to find in an assembly with heterogeneous voters? How and why do voting blocs change over time? These are some of the questions that I address in this paper, modeling an assembly with a finite number of agents who can coordinate with each other to form voting blocs before they vote to pass or reject a policy proposal.

I provide a new explanation for the emergence of political parties and the phenomenon of disciplined partisan voting. This explanation is based only on the properties of majority voting as a rule to aggregate preferences and it is therefore more broadly applicable than complementary theories that rely on campaigns and elections, on particular institutions that determine the voting agenda, or on any other element besides the act of voting in an assembly.

According to my theory, party discipline and whipped voting do not stem from existing political parties and their sophisticated partisan strategies. Rather, political parties are born from the gains to be made by forming a voting bloc. A group of members of an assembly -a party- strategically coalesce into a voting bloc to coordinate their votes, seeking to influence the policy outcome for an ideological gain. Party members commit to accept the party discipline and to vote for the party line, which is chosen according to an aggregation rule internal to the party.

I consider an assembly in which any subset of voters can coalesce to form a voting bloc. I analyze the endogenous formation of voting blocs and I show that in equilibrium, voting blocs form and exercise party discipline to affect voting records. To obtain sharper predictions about the voting blocs endogenously formed, I apply the model to a small assembly and I introduce a *split-proof* equilibrium refinement that allows for coalitional deviations in which at most one bloc splits apart. I show that in a stylized assembly with 9 members whose types are symmetrically distributed, two voting blocs form, one at each side of the ideological spectrum with a group of independents including the median in between the two blocs. In the last section of the paper I compare this result with the predictions derived from empirical data on the voting patterns of the United States Supreme Court from 1995 to 2004.

Using data on the 419 non-unanimous decisions that the Court reached in this period, I provide estimates of the ideal position of each justice in one and two dimensional spaces and I calculate how the formation of voting blocs would have changed the decisions of the Court. For each hypothetical set of connected voting blocs I find the decisions that would have been reversed due to the coordination of votes if these particular blocs had formed. I assume that the justices that dissented (voted with the minority) on a decision would have liked a reversal of the decision, and those who in the data voted with the majority and won would have been worse off had the decision been reversed. Aggregating over all the decisions, I calculate the net balance of beneficial minus detrimental reversals for each justice induced by the given voting blocs, relative to the original data. Assuming that these individual net balances of reversals are payoffs to the justices, I calculate the voting blocs that strategic justices would form in a split-proof equilibrium of the coalition-formation and voting game. The only split-proof equilibria involve the formation of two voting blocs of size three: Three of the four liberal justices (Stevens, Ginsburg, Souter, Breyer) in a liberal bloc, and the three most conservative justices, namely Rehnquist, Scalia and Thomas in a conservative voting bloc. This empirical exercise shows that justices have strategic incentives to coalesce into voting blocs.

1.1 Literature Review

In a companion paper,[1] I introduce a static version of the theoretical model, in which agents make only one policy decision. I first present a partial equilibrium analysis, where agents belonging to either of two pre-existing parties choose whether or not to accept party discipline and form voting blocs. I analyze the incentives of each agent to accept party discipline depending on the types of all agents, the size of the parties, and the voting rules that each party uses to aggregate internal preferences. I develop a fully endogenous theory of voting bloc formation in the second part of the paper, which presents several general theoretical results that hold under very weak assumptions about the existence of stable voting blocs for various definitions of stability.

In the current paper I extend the static model to a dynamic game where agents vote repeatedly, although I also show that under some stationarity assumptions, the static model serves as an adequate proxy for the play in the repeated game. I also present the theory of endogenous voting bloc formation in full game form, proposing a standard equilibrium concept first, and then a refinement to account for coalitional deviations. Finally, the last section of the current paper forgoes generality to seek instead sharp predictions in a given assembly, finding that in equilibrium, two polarized voting blocs form in a small assembly.

The theory in this paper and its companion draws inspiration from several literary subfields.

[1] Eguia, "Voting Blocs, Coalitions and Parties", revised May 2007. Available from the author or from the web at http://politics.as.nyu.edu/object/JonEguia.

Carraro [9] surveys recent non-cooperative theories of coalition formation, but mostly with economic and not political applications. Traditional models of coalition formation assumed that agents only care about by the coalition they belong to, not by the actions of other agents outside their coalition. In contrast, the *partition function* approach first used by Thrall and Lucas [35] recognizes that agents are affected by the actions of outsiders, and it defines utilities as a function of the whole coalition structure in the society. Bloch [6] and Yi [36] survey the literature on coalitions that generate positive externalities to non-members, such as pollution-control agreements, and coalitions that create negative externalities to non-members, such as custom unions. More recently, Bloch and Gomes [7] propose a general model to cover a variety of applications with either positive or negative externalities. However, there is no literature yet on the more general case in which a coalition generates both positive and negative externalities to non-members. A forthcoming paper by Hyndman and Ray [19] makes the first contribution to this future literature in a restricted model with only three agents. The formation of a voting bloc or a political party of any size generates positive externalities to those who agree with the policies endorsed by the party, and negative externalities to agents with an opposed policy preference. My model provides intuitive results for the mixed or hybrid case in which the formation of a voting bloc or party generates both positive and negative externalities to non-members, in a simple framework where the outcome of a voting game determines the payoff to each of finitely many agents.

In previous formal theories of party formation, Snyder and Ting [33] describe parties as informative labels that help voters to decide how to vote, Levy [23] stresses that parties act as commitment devices to offer a policy platform that no individual candidate could credibly stand for, Morelli [25] notes that parties serve as coordination devices for like-minded voters to avoid splitting their votes among several candidates of a similar inclination. All these theories explain party formation as a result of the interaction between candidates and voters. Baron [3] and Jackson and Moselle [20] note that members of a legislative body have incentives to form parties within the legislature, irrespective of the interaction with the voters, to allocate the pork available for distribution among only a subset of the legislators. My theory shows that legislators also have an incentive to form parties -voting blocs- in the absence of an electoral or distributive dimension, merely to influence the policy outcome over which they have an ideological preference.

In the American Politics literature, Cox and McCubbins [10] find that legislators in the majority party in the US Congress use the party as means to control the agenda and the committee assignments, and Aldrich [1] explains that US parties serve both to mobilize an electorate in favor of a candidate, and to coordinate a durable majority to reach a stable policy outcome avoiding the cycles created by shifting majorities. I complement their explanations proving that voting blocs of size less than minimal winning also influence the outcome even if they are not big enough to guarantee a majority, and they generate an ideological policy gain to their members.

A strand of the political economy literature studies the formation of governments by coalitions of parties. Four decades after Riker [28] showed the advantages of forming minimal winning coalitions, Diermeier and Merlo [11] show that if agents bargain over ideology and not just the distribution of resources, coalitions may be smaller or larger than minimal winning, and in a recent book, Schofield and Sened [30] survey the latest theoretical and empirical findings about the formation of government coalitions in multiparty democracies. Finally, the voting power literature exemplified by the work of Gelman [18] takes a different approach on coalition formation and assumes that agents want to maximize the probability of being pivotal in the decision, instead of maximizing the probability that the outcome is favorable to their interests.

In the following sections I apply the game-theoretic insights of the coalition formation literature to explain the coordination of votes by members of an assembly and to predict the formation of political parties. First I offer an intuition for the model using two examples.

1.2 Motivating Examples

In this subsection I present two examples to illustrate how the formation of voting blocs affects voting results and policy outcomes.

Example 1 *Let there be an assembly with five agents A, B, C, D, E who have to make a binary choice decision -to approve or reject some action- by simple majority. Suppose that only agents C, D favor the proposal, so if agents vote their individual preference, the proposal is rejected 2-3. If agents B, C, D form a voting bloc that commits to vote together according to the preferences of the majority of members of the bloc, C, D in favor of the action achieve an internal majority and with the three votes of the bloc B, C, D the action is approved. This outcome makes B worse off so in this case B has no incentives to join such a bloc. However, suppose instead that there are three different actions to be approved or rejected on three different topics. Suppose further that only C, D favor action one, only B, C favor action two and only B, D favor action three. Then, without voting blocs, all actions are rejected 2-3. Agents B, C and D get their desired outcome only in one decision. If they form a voting bloc, the majority of the bloc favors all three actions, and all three are approved. Agents B, C and D are all better off, achieving their desired outcome in two decisions.*

In the example, each member of the voting bloc benefits from joining in. Each agent is forced to vote against her wishes in one instance, but more often (twice), belonging to the bloc allows the agent to gather enough votes to sway the decision of the assembly to her preferred outcome. The same incentive to join a voting bloc exists if agents vote over a single decision, but with some uncertainty over preferences. For instance, suppose in the above example that there is only one decision to make, but initially it is uncertain whether B, C, or B, D, or C, D will be the two agents in favor of the action while all others are against it. Then the incentives of B, C, D to form a bloc are the same: Without a bloc their probability of achieving their desired outcome is one third, with a bloc it is two thirds.

In the model that I present, I consider both uncertainty and repeated play, so that agents with some uncertainty over preferences vote on a sequence of policy proposals. We can interpret the uncertainty about preferences in two complementary ways. First, suppose there is a time difference between the moment when agents coalesce in voting blocs, and the time of voting in the assembly. Then, when the agents make the commitment to act together they do not fully know which outcome they will prefer at the time of voting. Three legislators may sign a pact today to vote together in votes to come in the future, but they do not know today the agenda or the details of the policies they will vote on in the future. Alternatively, in a world in which agents vote repeatedly, a legislator who votes for the liberal policy with a certain frequency x can be modeled as a legislator with a probability x of voting for the liberal policy each time.

The voting power literature[2] focuses on an extreme case of uncertainty, where agents not only do not know exactly how they will feel about future policy proposals, but they can't even take a guess. In my paper, I assume that there is some uncertainty about how agents vote, but that ex-ante it is possible to differentiate agents according to their expected preferences. For instance, it is not a foregone conclusion that a Republican legislator in the US Senate will vote in favor of future tax cuts and a Democratic senator against them, but it is ex-ante more probable that the Republican, rather than the Democrat, will favor the tax cuts.

The ex-ante differences in the preferences of the agents are key determinants of the strategic incentives to form voting blocs. Intuitively, agents prefer to coalesce with other like-minded voters.

Example 2 *Let there be an assembly with nine agents who must make a binary choice decision -pass or reject some policy proposal- by simple majority. Suppose that agents have uncertain preferences, so that each agent i favors the proposal with an independent probability w_i. Suppose $w_l = 0.15$ for $l = 1, 2, 3, 4$, $w_5 = 0.5$ and $w_h = 0.85$ for $h = 6, 7, 8, 9$. Table 1 shows the probability that the outcome coincides with the preference of a given agent, expressed as a percentage, given that the following voting blocs form: no blocs (row one); agents $1, 2, 3$ form a bloc (row two); agents $1, 2, 3, 4$ form a bloc (row three); agents $1, 2, 3, 5$ form a bloc (row four); and agents $1, 2, 3, 6$ form another bloc (row five). If a bloc forms, the whole bloc votes according to the preference of the majority of its members, and in case of a tie, each member votes according to her own preferences.*

Bloc	1	2	3	4	5	6	7	8	9
None	59.0	59.0	59.0	59.0	70.9	59.0	59.0	59.0	59.0
$\{1, 2, 3\}$	63.6	63.6	63.6	69.9	73.5	48.8	48.8	48.8	48.8
$\{1, 2, 3, 4\}$				67.2					
$\{1, 2, 3, 5\}$					63.7				
$\{1, 2, 3, 6\}$						43.4			

Table 1: Probability that agents get their desired outcome, in %.

[2] Within this literature, see Felsenthal and Machover [16] for a study of voting blocs.

The numbers on the table come from simple binomial calculations. Note that the formation of a voting bloc by agents $1, 2, 3$ has a significant effect on the outcome, even though this bloc does not command in itself a majority, unlike the bloc in the simplistic Example 1. The three agents that form the bloc increase their probability of achieving their desired policy outcome by four percentage points, so none of them have an incentive to abandon the bloc and disband it. The last three rows of the table show that no other agent has an incentive to join the voting bloc, and therefore, the formation of this bloc is a Nash equilibrium. Surely, it is not a unique Nash equilibrium: The same calculations apply if $2, 3, 4$ form a voting bloc instead, or $6, 7, 8$ among other possibilities. I propose solutions for this multiplicity below, merely noting by this example that there exist Nash equilibria in which agents form voting blocs to coordinate their votes in such a way that they affect policy outcomes to their benefit.

The insights gained in these two examples have apply to voting in committees, councils, assemblies, and, in particular, in legislatures where legislators can coalesce into political parties that function as voting blocs.

2 The Model

Let N be an assembly of voters $i = 1, ..., n$, where n is odd. This assembly chooses the policy outcome in each of finitely many stages $t = 1, ..., T$. In each stage, a policy proposal is exogenously given, and the assembly makes a binary decision on whether to adopt this proposal, or reject it, in which case a default policy is implemented. Slightly abusing notation, let the policy proposal put to a vote in stage t be labeled proposal t. At each stage, the assembly chooses the policy outcome by majority voting. The division of the assembly is the partition of the assembly into two sets: Those who vote in favor of proposal t, and those who vote against of t. If the number of votes in favor is at least nr_N, then the proposal passes, otherwise proposal t fails and the default policy is implemented at this stage. In either case, policy proposal $t+1$ is put to a vote in the following stage.

At stage t, voter $i \in N$ receives utility one if the policy outcome coincides with her preference in favor or against proposal t and zero otherwise. Utility is additive over time with no discounting. Since agents have neither varying intensity over preferences, nor discounting over time, their optimization problem is to maximize the number of stages in which the policy outcome coincides with their binary preference for or against the proposal. Let $p_i^t = 1$ if agent i prefers proposal t to pass, and zero otherwise; let $p^t = (p_1^t, ...p_n^t)$ indicate the profile of preferences at stage t of the whole set of voters, and let $p_{-i}^t = (p_1, ..., p_{i-1}, p_{i+1}, ..., p_n)$ be the profile without the preference of i. Similarly, let $v_i^t = 1$ if agent i votes in favor of proposal t in the division of the assembly, and $v_i^t = 0$ otherwise. Then, policy proposal t passes if and only if $\sum_{i \in N} v_i^t \geq nr_N$, where r_N is the voting rule used by the assembly.

Agents face uncertainty at the beginning of each stage. They do not know the profile of preferences in favor or against the proposal. They only know, for each profile of preferences $p^t \in P = \{0, 1\}^n$, the probability that p^t occurs. Let $\Omega^t : \{0, 1\}^n \longrightarrow [0, 1]$ be the probability distribution over preference profiles at stage t and assume that Ω^t is common knowledge at the beginning of stage t.[3] This uncertainty is resolved and agents privately learn their own preference before the policy proposal comes to a vote. However, prior to the resolution of the uncertainty about preferences, and knowing only Ω^t, agents can coalesce into voting blocs.

Any subset of the assembly $C_j \subset N$ can coordinate the voting behavior of its members by forming a voting bloc $V_j = (C_j, r_j)$ with an internal voting rule r_j that maps the preferences of its members into votes cast by the bloc in the division of the assembly. Then it becomes a voting bloc. I assume that joining a voting bloc is voluntary and agents may also remain independent. Formally, I assume that there exists a list or vector of rules $\mathbf{r} = (r_0, r_1, ..., r_R)$. These rules function as contracts. Agents choose which contract to sign, and all agents that sign the same contract belong to the same voting bloc. Contracts are binding for only one stage, they do not commit agents in any way for future stages. Each agent must sign exactly one contract, but rule r_0 specifies that all signers remain independent. All other rules specify a commitment on the part of the signers to coordinate their votes in their assembly as detailed below.

If a coalition C_j of size N_j forms a voting bloc with rule r_j at stage t, in an internal meeting prior to the assembly meeting, the members of C_j vote to determine their coordinated behavior in the division of the assembly according to their own internal rule r_j and I assume that the voting bloc has commitment mechanisms such that the outcome of this internal meeting is binding for the vote in the division of the assembly at stage t. In particular, each member of C_j casts an internal vote $\widehat{p}_i^t \in \{1, 0\}$ for or against the proposal, and these internal votes are aggregated into a common outcome for the bloc with three possibilities:

1. If $\sum_{i \in C_j} \widehat{p}_i^t \geq r_j N_j$, then $\sum_{i \in C_j} v_i^t = N_j$. If the fraction of C_j members who favor the policy proposal is at least r_j, then the whole bloc votes for the proposal in the division of the assembly.

2. If $\sum_{i \in C_j} \widehat{p}_i^t \leq (1 - r_j) N_j$, then $\sum_{i \in C_j} v_i^t = 0$. If the fraction of C_j members who are against the policy proposal is at least r_j, then the whole bloc votes against the proposal in the division of the assembly.

3. If $(1 - r_j) N_j < \sum_{i \in C_j} \widehat{p}_i^t < r_j N_j$, then $\sum_{i \in C_j} v_i^t = \sum_{i \in C_j} \widehat{p}_i^t$. If neither side gains a sufficient majority within the voting bloc, the bloc does not act together and it reproduces in the division of the assembly the same internal split.[4]

The timing of stage t is as follows:

1. The probability distribution over preferences Ω^t and the list of available rules $\mathbf{r}$ are public knowledge. Agent i remains uncertain about the exact preference p_i^t.

[3] We can assume either that at the beginning of stage one Ω^t is public knowledge for every t, or alternatively, that Ω^t follows some exogenously given stochastic process over time which is common knowledge. The relevant assumption is that Ω^t is exogenous and public knowledge at the beginning of stage t.

[4] The results below don't change if in this third case we allow agents to change their vote from the internal meeting to the meeting of the assembly because, as I will show, the revealed preferences $\widehat{p}_i^t$ coincide with the sincere preferences p_i^t in the equilibria of interest.

2. Each agent i chooses a rule from the list $\mathbf{r}$. The set of agents C_j who choose a given rule r_j form a voting bloc $V_j = (C_j, r_j)$.

3. Each agent i privately learns p_i^t, her preference for or against policy t.

4. Voting blocs meet. Each member of a bloc casts a vote for or against the proposal, and these votes, together with the internal rule of the bloc, determine the actions of the bloc in the division of the assembly.

5. The assembly meets. Agents vote according to the outcome of substage 4, and the proposal passes if it gathers enough favorable votes.

The intuition of this timing is that initially agents may have only a rough idea about the policy proposal. Perhaps they know that they will debate a new public health-care bill. Agents know who is likely to favor or oppose a public health-care bill, but they can't be sure since they have not read the details of the bill yet. At this point, agents form alliances and coalesce into groups, committing to discuss the bill internally and act as a bloc once the bill comes to the floor. Then a lower chamber, or a committee, or an arbitrary exogenous body, produces the bill for agents to inspect it, and voters learn their true preference. Voting blocs then meet to aggregate the preferences of their members, revealed by the votes at the internal meeting. Finally the assembly meets after blocs have committed to a coordinated voting strategy and the bill passes if it gathers enough votes.

I model agents who choose to remain independent as joining a fictitious voting bloc V_0 with rule $r_0 = 1$. With this unanimity rule, the vote that the agent freely casts at the internal meeting always coincides with the final vote v_i^t in the division of the assembly, so these independent agents do not coordinate their votes and vote as they wish in the assembly. In all other blocs, a sufficiently high majority rolls an internal minority, forcing the minority to vote with the majority of the bloc in the division of the assembly.

I assume that the list of available rules contains enough rules so that, for any size N_j and any integer x strictly larger than $\frac{N_j}{2}$ and strictly smaller than N_j there exists r_j such that $x \leq r_j N_j < x + 1$. That is, voting blocs of any size have contracts available to them setting an internal threshold to coordinate votes at any desired number greater than half the size of the bloc. Or, in other words, any majority rule, from simple majority, to all-but-one supermajority is available to blocs of any size. Furthermore, there are enough copies of each voting rule or contract so that several blocs can form, all of them with an identical internal rule. For instance, any two disjoint coalitions C_j and $C_{j'}$ can form a separate voting bloc with simple majority internal rule by choosing $r_j = r_{j'} = \frac{n+1}{2n}$, regardless of how many other coalitions are also forming voting blocs with simple majority.

Let h^t be the history of the game up to stage t. This history includes the probability distributions over preferences of all previous periods, $(\Omega^\tau)_{\tau=1}^{t-1}$ as well as the actions of all agents in all previous periods.

Let s_i^t denote the pure stage strategy of agent i at stage t, which specifies two elements. First, given the history h^t and the current probability distribution over preference profiles Ω^t, s_i^t determines the contract a_i^t that i signs, and s_i^t also determines the vote $\widehat{p}_i^t$ of agent i for or against proposal t inside the voting bloc, as a function of the true preference of i and the contracts chosen by all agents $a^t = (a_1^t, ..., a_n^t)$. Formally,

$$s_i^t\left(h^t, \Omega^t\right) = \left(a_i^t, \widehat{p}_i^t(p_i^t, a^t)\right) \text{ and } h^t = (\Omega^\tau, a^\tau, \widehat{p}^\tau)_{\tau=1}^{t-1}.$$

A pure strategy for the game for agent i is a finite sequence of T pure stage strategy functions, one for each stage: $s_i = (s_i^t\left(h^t, \Omega^t\right))_{t=1}^T$. A pure strategy profile for the game is then $s \in S$, where $s = (s_1, ..., s_n)$ and S is the set of all feasible strategies profiles. I denote by $s^t = (s_1^t, ..., s_n^t)$ the pure strategy profile at stage t, so note that s is also the finite sequence of stage strategy profiles $s = (s^t)_{t=1}^T$.

The stage utility of agent i, defined at the beginning of stage t, is the ex-ante probability that the policy outcome at the end of the stage coincides with the preference of the agent, given the probability distribution over preferences, and the stage strategies of every agent. I denote this utility by $u_i^t(s)$. No discounting means that the aggregate ex-ante utility of agent i, evaluated at the beginning of the first period, is equal to $\sum_{t=1}^{T} u_i^t(s)$. The goal of each agent is to maximize this aggregate utility, by choosing which voting blocs to join, and how to vote. Note that Ω^t determines the expected payoffs of the stage game t as a function of the actions of the players. Let $\Omega = (\Omega^t)_{t=1}^T$ be the sequence of probability distributions over preferences at each stage, which in turn determines the payoffs of the whole game.

Then $\Gamma = (N, S, \Omega)$ denotes the political game of coalition formation and voting that the n agents play in T stages.

A subgame perfect Nash equilibrium of the game Γ specifies strategies for all agents that are mutual best responses for any subgame of the game. Given the multiplicity of such equilibria with uninteresting properties (such as a trivial equilibrium where no one joins a bloc and everyone votes against every proposal, so no proposal passes and no unilateral deviation can make an agent better off), I look for equilibria that satisfy two added properties: stationarity, and weak stage-dominance.

Definition 1 *A strategy profile s is stationary if the stage strategy s_i^t is independent of history h^t for all $t \in T$ and all $i \in N$.*

The probability distribution over preferences Ω^t determines the expected payoffs of the stage game. Stationarity as defined means that faced with a given stage game at a given time, an agent uses the same stage strategy regardless of what happened at any previous stage. Stationarity rules out inter-temporal punishment of the signing or voting decisions made by an agent at any previous time. Consequently, in stationary equilibria the agents are able to maximize their stage utility myopically and at the same time maximize the aggregate inter-temporal utility, which reduces the complexity of the optimization problem.

Stationarity alone does not prevent implausible equilibria in which every agent votes against the proposal, so no agent has an incentive to deviate. In a one-shot game, such equilibria are discarded assuming that agents never play weakly dominated strategies. Following Baron and Kalai [4] and Duggan and Fey [12], I extend this notion of undomination assuming that voters, while holding the strategies of all players in future stages as fixed, do not play stage strategies that are weakly dominated. At any voting stage, and considering the stage game in isolation, voting for the least preferred of the two policy alternatives is weakly dominated. *Stage undomination* requires that voters do not vote for their least preferred candidate at any stage unless they expect to gain something from having cast such a vote in a future stage.

Definition 2 *Let $S_s \subset S$ be the set of all strategy profiles in which exactly one insincere vote from strategy profile s is reversed and made sincere. The strategy profile s is stage undominated if $\sum_{\tau=t+1}^{T} u_i^\tau(s) > \sum_{\tau=t+1}^{T} u_i^\tau(s')$ for any strategy profile $s' \in S_s$ such that the vote reversed to sincerity is cast by agent i at stage t.*

An insincere vote cannot yield a higher present stage payoff. Stage undomination only states that such a vote is not cast unless it yields a higher aggregate payoff in future stages, given the existing strategies. Stage undomination together with stationarity imply sincere voting. Since there is no obvious notion of sincerity in the game of coalition formation, I say that a strategy profile is sincere if voting at the internal meetings is sincere.

Definition 3 *A strategy profile s is sincere if $\widehat{p}_i^t(p_i^t, a^t) = p_i^t \; \forall i \in N, \forall t \in T$.*

Lemma 1 *If a strategy profile s is stationary and stage undominated, then it is sincere.*

The proof is immediate: If a strategy profile is stationary, stage strategies and stage payoffs in all future stages are, by definition, independent of the vote that an agent casts in the current stage. Since insincere voting is weakly dominated in the stage game, if the agent casts an insincere vote in the current stage, it must then do so in violation of stage undomination.

The equilibrium concept I use is Stage-Undominated Stationary Pure-Strategy Subgame-Perfect Nash Equilibrium, from now on referred simply as an *equilibrium.* While it is possible to find equilibria in which voting in the assembly is unaffected by voting blocs -either because these don't form, or if they do their members all agree in their preferences so no internal rolling takes place,- the more interesting case is the existence of equilibria in which voting blocs form and affect voting records by aggregating internal preferences in such way that exercises party discipline, so that agents cast votes in the assembly against their individual preferences, as dictated by the rules of the bloc they belong to.

Definition 4 *A sincere strategy profile s exhibits party discipline if it is such that, with positive probability, $v_i^t \neq p_i^t$ for some agent i and stage t.*

In a strategy profile with sincere voting, agents vote for their preferred alternative in the internal meeting of their voting bloc and, if they are independent, in the assembly. An agent who follows a sincere voting strategy only casts a vote against her preference if she loses the internal vote of her bloc, and must then deliver her vote to the majority of her bloc in the division of the assembly.

Theorem 2 *Suppose $r_N \leq \frac{n-1}{n}$ and suppose Ω^t has full support for some t. Then there exists an equilibrium with party discipline.*

Proof. Let s be such that $s_i^t(h^t, \Omega^t) = (a_i^t, \widehat{p}_i^t(p_i^t, a^t)) = (r_1, p_i^t)$ with $r_1 = \frac{n+1}{2n}, \forall i \in N, \forall t \in T$ and for any history h^t. I first show that s exhibits party discipline. Take t such that Ω^t has full support, then $p^t = (1, 0, 0, ..., 0)$ with positive probability. Agent 1 loses the internal vote, and $p_1^t = 1$ but $v_1^t = 0$. Second, I note that the proposed strategy is stage undominated, since no agent ever votes for her least preferred alternative; stationary, since the strategy is history-independent; and pure. I now show that no deviation can make an agent i better off at any subgame. First, deviating at stage t by choosing $a_i^t \neq r_1$ or $\widehat{p}_i^t \neq p_i^t$ has no effect on the play or payoffs on any other stage, since s is stationary. Hence it suffices to analyze the incentives to deviate in each stage game. If an agent deviates by voting insincerely, either the agent is not pivotal and the deviation has no effect, or the agent is pivotal and her deviation changes the outcome and lowers her payoff. Finally, suppose the agent deviates by choosing $a_i^t \neq r_1$. Note that with a rule in the assembly which is not unanimity, the voting bloc acts as a dictator, even after the defection of one agent. Therefore, if the defecting agent votes against the preferences of the majority of the bloc, she loses and attains utility zero at this stage; if the defecting agent votes with the preferences of at least one half of the remaining members of the bloc, then she wins, but she would also win belonging to the bloc, so she gains nothing by deviating. In either case, agent i has no incentives to deviate at stage t, for any t or any combination of stages. ■

In the following section of the paper I sharpen this existence result, making specific predictions about the equilibria that arise in a theoretical assembly with stylized preferences, then with preferences inferred from the roll-call votes in the US Senate. First I present a second general result, which is important to interpret the endogenous voting blocs that emerge in each stage as permanent, stable political parties. Let $\Gamma^1 = (N, S, \Omega^t)$ denote the one-shot game with preferences given by Ω^t. This is stage game t considered in isolation, disregarding history and future stages. In a repeated finite game, the strategy consisting of playing a Nash equilibrium of the stage game at each stage is a Nash equilibrium of the repeated game. This is a fairly well-known result in repeated game theory, noted for instance by Fudenberg and Tirole [17], p 149. Applying this argument to my model, I find that maintaining the same voting blocs over time is an equilibrium strategy as long as the preferences that sustain such voting blocs do not change.

Theorem 3 *Suppose $\Omega^{t+\tau} = \Omega^t$ for any $\tau = 1, 2, ...k$. Let the strategy profile s^t be an equilibrium of the stage game $\Gamma^1 = (N, S, \Omega^t)$. Then the strategy profile s consisting of playing s^t at any stage after any history is an equilibrium of the repeated game with $k+1$ stages beginning at t and ending at $t+k$.*

Proof. By induction. Consider first the subgame starting at stage $t+k$. By assumption, playing s^t at this stage is an equilibrium of this subgame, regardless of history. Now suppose that s^t is played in the equilibrium of the whole game at every stage from $t+\lambda$ to $t+k$. It suffices to show that s^t is then played at stage $t+\lambda-1$ to complete the induction argument. Consider the incentives of agent i to deviate at any stage $t' \in [t+\lambda-1, t+k]$. Since s^t is an equilibrium of the stage game t', agent i cannot achieve a higher stage payoff in t'. Since the stage strategies of all agents in stages after t' are determined by s^t, and this strategy indicates the same actions for all agents following the deviation by i, the payoff of agent i in all stages following her deviation are invariant. Hence she cannot gain at any stage, present or future, by deviating at stage t', or at any combination of stages (see Fudenberg and Tirole [17], Theorem 4.1. for a proof that in a finite game with observable actions, if agents cannot gain from deviating in just one stage, then they cannot gain by deviating at several stages). Since no agent has an incentive to deviate at stage $t+\lambda-1$, s^t is played at this stage and the induction argument is complete. ■

As long as the priors over preferences of the members of the assembly given by Ω^t do not change, it remains an equilibrium of the stage game for the same voting blocs to persist over time. I interpret these long-lasting, permanent voting blocs that exercise party discipline as political parties, which only break up when a change in preferences makes the current division into parties unstable. Note that this permanence of the same voting blocs in equilibrium is not trivially implied by stationarity. Stationarity requires that at a given stage, the stage formation of blocs must be the same for any history as long as the probability distribution over preferences is the same, but it does not rule out a different configuration of voting blocs at each stage. A bit more formally, stationary stage strategies depend on the pair (t, Ω^t). Theorem 3 shows that a stronger form of stationarity holds, in which stage strategies depend only on Ω^t and not on the stage t at which Ω^t occurs.

3 Endogenous Voting Blocs in a Small Assembly

3.1 Two Polarized Blocs Emerge in a Stylized Assembly

Consider an assembly with nine voters and a simple majority voting rule $r_N = \frac{5}{9}$. Suppose that agents have independent preferences. That is, Ω^t is uncorrelated, each agent i has a prior w_i^t, which is the probability that i favors proposal t, and the preference profile p^t occurs with probability $\prod_{i=1}^{9} w_i^t p_i^t + (1-w_i^t)(1-p_i^t)$. Equivalently, $\Pr[p_i^t = 1 | p_{-i}^t] = w_i^t$ for all i, t and p_{-i}^t.

Suppose further that Ω^t is fixed over time and the distribution of priors is symmetric as follows: w_1=w_2=0.5-α-β; w_3=w_4=0.5-α; w_5=0.5; w_6=w_7=0.5+α; w_8=w_9=0.5+α+β, with $\alpha, \beta \geq 0$ and $\alpha + \beta \leq 0.5$. In words, priors are symmetrically distributed around one half.

The parameters α and β have an intuitive interpretation: α measures the polarization of preferences within the assembly. A hypothetical coalition of moderates comprising agents 3 through 7 (enough to become a majority centered around the median) spans an interval of priors of length 2α. A more polarized assembly corresponds to a larger α and larger differences in priors that a coalition of moderates must accommodate in order to form a voting bloc. The parameter β, albeit crudely, reflects the heterogeneity in types within each side of the assembly, or in other words, the extremism of the left-most and right-most wings.

An intuitive conjecture is that intense polarization in the assembly would make a central voting bloc unstable and would induce the formation of two opposing voting blocs, one on each side of the median.

To formulate a theoretical prediction using the model, given that Ω^t is fixed over time, I take advantage of Theorem 3 to look only at the equilibrium of a simplified one-stage game, constructing the equilibrium of the whole game as the repetition of the stage equilibrium at all stages. I can then drop the superindex t from all variables without confusion.

Theorem 2 informs us that equilibria with party discipline exist. There often exists multiple equilibria, but some are more interesting than others. In particular, I seek equilibria such that party discipline not only affects voting records by rolling a few isolated, non-pivotal votes, but rather, the formation of voting blocs affects policy outcome, by making the majority of votes in the division of the assembly not correspond with the majority of sincere preferences. In these equilibria, party discipline is relevant for the policy outcome.

Definition 5 *A sincere strategy profile s exhibits relevant party discipline if either* $\left(\sum_{i\in N} p_i < nr_N \leq \sum_{i\in N} v_i\right)$ *or* $\left(\sum_{i\in N} v_i < nr_N \leq \sum_{i\in N} p_i\right)$ *occurs with positive probability.*

In the specific assembly with nine agents and simple majority, a strategy profile includes relevant party discipline if the aggregation of votes inside a bloc makes the set of agents with the minoritarian preference in the assembly gather five votes in division of the assembly, by means of achieving a sufficient majority in at least one voting bloc, and rolling the votes of the defeated members of the bloc.

Example 1 in the Introduction considered an assembly with $\alpha = 0.35$ and $\beta = 0$. Note that there are several equilibria with relevant party discipline in this assembly: Agents $1, 2, 3$ may form a voting bloc, or agents $2, 3, 4$, or instead agents $6, 7, 8$ or $7, 8, 9$ can form a unique voting bloc in equilibrium. If any of these blocs form, no agent has an individual incentive to deviate. However, all these equilibria are intuitively unsatisfactory. If the agents with a low prior form a voting bloc to their advantage and to the detriment of the agents with a high prior, it begs the question: why don't agents with a high prior form their own voting bloc as well? The game-theorist response is that under Nash equilibria, agents can only deviate unilaterally, taking the strategies of other agents as given. Table 2 adds one more row to Table 1, to compare the outcome if both the agents with a low prior and the agents with a high prior form a voting bloc

Blocs	1	2	3	4	5	6	7	8	9
{123}	63.6	63.6	63.6	69.9	73.5	48.8	48.8	48.8	48.8
{123}, {789}							53.3	53.3	53.3

Table 2: Table 1 extended to consider two voting blocs.

If three high-prior agents form a voting bloc in response to the low-prior voting bloc, they increase the probability that the policy outcome coincides with their individual preference from a bit less than one-half, to a bit more than one-half. But under Nash equilibria, agents best-respond unilaterally, and agents $7, 8, 9$ cannot coordinate a coalitional deviation away from the equilibrium with only one bloc, even though the outcome with two blocs is itself an equilibrium. Hard as it may be for agents to communicate and coordinate across preexisting blocs, it seems easier to scheme deviations involving only independents or agents of a single bloc, or a mix of both defecting together and possibly forming a new voting bloc. The following equilibrium concept allows for a coalitional deviation in which one bloc faces a split, a number (possibly zero) of its members defect, and at the same time a (possibly empty) subset of the defectors and previously independent agents form a new voting bloc.

Definition 6 *An equilibrium strategy profile s is split-proof if there exist no set of agents $E \subseteq N$, rule r_j and strategy profile $s' \in S$ such that:*

(i) $a_i \in \{r_0, r_j\}$ for all $i \in E$,

(ii) $a'_i \in \{r_0, r_{j'}\}$ with $r_{j'}$ such that for any $l \in N$, $a_l \neq r_{j'}$,

(iii) $s_h = s'_h$ for any $h \notin E$, and

(iv) for any $i \in E$ s.t. $a'_i = r_0$, $u_i(s') \geq u_i(s)$ and for any i s.t. $a'_i = r_{j'}$, $u_i(s') > u_i(s)$.

Condition *(i)* says that all the agents who coordinate a deviation are initially either members of the bloc with rule r_j, or independents. Condition *(ii)* states that after the deviation, all the deviants become either independents, or members of a new bloc with a rule $r_{j'}$ that previously had attracted no membership. Condition *(iii)* states that the rest of the agents do not react to the deviation in any way. Condition *(iv)* states that agents who defect become better off. When agents are indifferent between deviating or not, this fourth condition incorporates an intuitive discrimination: Agents may abandon a bloc to become independents when indifferent, but they only deviate to a new bloc for a strict improvement. That is, agents break indifference as if they had a lexicographic preference for independence.

The intuition for the split-proof equilibrium is that coalitional deviations across blocs are harder to coordinate, perhaps because communication is limited across blocs, or because different blocs are antagonistic and suspicious of each other (i.e. Western and Soviet blocs during the Cold War); whereas, a disaffected subset of a bloc can more easily break apart and possibly recruit some independent agents for a new voting bloc. As an example, the moderate wing of the UK's Labour party broke off in 1981 and formed the Social Democratic Party, which attracted up to 28 former Labour MPs.[5]

The notion that some members of a coalition may organize a coordinated defection even though deviations across coalitions are not feasible is common to two previous concepts of equilibrium in the non-cooperative coalition formation literature: The Coalition-Proof equilibrium by Bernheim, Peleg and Whinston [5] and the Equilibrium Binding Agreements by Ray and

[5] Admittedly, cross-party deviations are sometimes also successful, as illustrated by the new Kadima party in the Israeli Knesset.

Vohra [27]. In these two concepts, agents negotiate as if each coalition was in a separate room, and any group of agents in the same room could leave and find a new room for themselves, with the important proviso that every deviation must itself be immune to further deviations (once they deviants reach their new room, it must be that no subset of them would want to leave for yet another room), so the definitions are recursive.

A split-proof equilibrium is different first in that it is not a recursive concept, since I don't require a coalition of deviants to be immune to further deviations. Second, while I do not consider deviations across coalitions, I allow deviants to coordinate with independents. Under split-proof equilibria, agents negotiate as if each coalition was in its own room, but the independents were all in a central lobby, so that when a set of deviants departs from a coalition they can recruit any number of independents in their way to a new room.

The split-proof equilibrium follows more closely the *split stability* notion introduced by Kaminski [21] and [22] in an applied study of political parties in Poland. Parties satisfy *split stability* if they have no incentives to dissolve into smaller units, where the incentives are considered non-recursively. The split equilibrium in this paper requires not only that members of a party have no incentives to leave the party, but also that no subset of independents has an incentive to deviate and form a new voting bloc, either by themselves, or attracting defectors from one existing party.

In this section I find split-proof equilibria with connected voting blocs and relevant party discipline.

Definition 7 *A voting bloc V_j is* ***connected*** *with respect to the order $<$ if for all $h, i, k \in N$, ($a_h = a_k = r_j$ and $h < i < k$) implies $a_i = r_j$.*

Intuitively, the voting bloc V_j with rule r_j is connected if given any pair of agents who choose rule r_j, any other intermediate agent located between the original pair also chooses rule r_j. The order I use to define connectedness is according to the priors over preferences, w_i. A voting bloc is connected if its members are in consecutive positions in the ordering by priors. Axelrod [2] provides a detailed argument in favor of connected coalitions over non-connected ones.

Using numerical simulation for a fine grid of values of α and β, I find the solutions to the game. I assume that the list of rules available for agents to form voting blocs includes r_0 for independent agents who do not coordinate their votes, and at least three copies of rules $\frac{2}{3}, \frac{3}{4}, \frac{3}{5}, \frac{4}{5}, \frac{5}{6}, \frac{4}{7}, \frac{5}{7}, \frac{6}{7}, \frac{5}{8}, \frac{7}{8}, \frac{5}{9}, \frac{7}{9}$ and $\frac{8}{9}$, so that agents can form blocs of any size with any desired supermajority rule. The rule $\frac{5}{9}$ corresponds to simple majority for any bloc of any size.

The intuition that in a very polarized assembly there is no unique moderate bloc, but rather, two blocs one in each side of the political spectrum is verified. There are only four split-proof equilibria with connected voting blocs and relevant party discipline. These are all such that exactly two blocs $V_L = (C_L, r_L)$ and $V_R = (C_R, r_R)$ form, both of them with simple majority internal voting rules, each with three members, and $C_L \subset \{1, 2, 3, 4\}$, $C_R \subset \{6, 7, 8, 9\}$. That is, three of the four members of the assembly with a low prior choose to form a voting bloc with

Figure 1: a,b,c. Split-proof equilibrium voting blocs.

simple majority, and three members with a high type form another bloc. It is easy to visualize the V_L bloc as a “left” or pro-status quo party, which tends to vote against the policy proposal, and the V_R bloc as a “right” or reform party, which tends to vote for the policy proposal.

Figure 1a,b,c shows in black the parameter values for which each of the four outcomes is a split-proof equilibrium. For any $\alpha < 0.5$, the voting blocs exercise relevant party discipline.

The three figures share the common characteristic that the equilibria with relevant party discipline holds only for a high α. If the assembly is not polarized and agents share similar priors, then each agent in voting bloc V_L has an incentive to defect to V_R, effectively disbanding V_L since no bloc can function with only two members. If there is enough polarization, defections across blocs no longer occur.

As a summary, this subsection has shown that if the assembly is sufficiently polarized, there is a split-proof equilibrium with connected voting blocs and relevant party discipline. In this equilibrium, two opposing blocs form, one at each side of the median. In the rest of the section I depart from the stylized assumptions of the modelled assembly (symmetry and independence of priors), looking instead at real data from the United States Supreme Court. After introducing the Court and the policy preferences of its members, I calculate the effect of voting blocs upon the outcome of the Court.

3.2 The United States Supreme Court

The United States Supreme Court is the ultimate appellate court in the United States judicial system, and the arbiter of the United States Constitution. It is composed of nine justices and it uses a simple majority rule, so that the vote of five justices are enough to decide a case. The

Court makes a binary decision on the merits of each case: It either affirms or reverses the ruling of a lower court. In an accompanying Opinion, the Court provides the argumentation for its decision, and this Opinion serves as precedent for future cases.

I use the data on the decisions of the Court from *The United States Supreme Court Judicial Database* compiled by Spaeth [34] and I select all non-unanimous cases with written opinions in which all nine justices participate.[6] Spaeth codes the votes of each justice as zero or one depending on whether the vote to affirm or reverse the decision of the previous court is interpreted as more liberal or more conservative. An alternative binary coding of the votes which is unambiguously objective divides the votes between votes with the majority, and dissents -votes with the minority.

Table 3 shows the number of liberal votes and the number of dissents that each justice cast in the 419 non-unanimous decisions recorded from 1995 and 2004. The nine justices, abbreviated by the first three letters of their surname, are: Stevens, Ginsburg, Souter, Breyer, O'Connor, Kennedy, Rehnquist, Scalia and Thomas.

	1.Ste	2.Gin	3.Sou	4.Bre	5.O'Co	6.Ken	7.Reh	8.Sca	9.Tho
Liberal	344	308	307	276	160	155	98	84	71
Dissent	203	159	136	141	71	78	115	161	156

Table 3: Liberal and dissenting votes in 419 decisions.

The most extreme justices, either liberal or conservative find themselves in the minority of dissenters more often than the moderate justices. Justice O'Connor, traditionally regarded as the swing justice, dissents in only about one in six cases, while Justice Stevens, who is the most liberal member of the Court, dissents from the majority in roughly a half of the cases.

If justices formed voting blocs, the coordination of votes would change the voting record of the justices, the composition of the majority and dissent justices in each case, the outcome of some decisions, and, assuming that justices are policy-oriented, the utility or satisfaction of the justices with the outcome of the Court. I calculate the changes brought by the formation of any connected voting blocs in the Court.

The notion of a connected voting bloc requires an ordering of justices from one to nine. In the tables and the text I of this section I use the ordering according to the number of liberal votes cast as recorded by Spaeth [34]. I check if this ordering is robust by means of calculating the ideal location of the justices in a space vector using three mathematical methods that abstract from the substantive content of each case and attend only to the voting patterns and correlations across the justices. Although my basic goal is to obtain an objective and robust ordering of the justices, these analyses have an intrinsic value in that they provide estimates of the location of each justice in a vector space with an easy interpretation in ideological terms

[6] The unit of analysis in my data is the case citation (ANALU=0), the type of decision (DEC_TYPE) equals 1 (orally argued cases with signed opinions), 6 (orally argued *per curiam* cases) or 7 (judgments of the court), and I drop all unanimous cases and all cases in which less than 9 justices participate in the decision.

such as a liberal/conservative scale.

The three methods I use are: Singular Value Decomposition of the original data, Eigen Decomposition of the square matrix of cross-products of the locations of the justices, and the Optimal Classification method developed by Poole [26], and I compare these three estimates with the findings of Martin and Quinn [24] and [15], who use Bayesian inference in a probabilistic voting to estimate the ideal points of the justices.[7] In Table 4 I provide the ideal position of the justices estimated by Single Value Decomposition and Eigen Decomposition, the rank ordering given by the Optimal Classification method in one dimension, the estimate of the position in the first dimension given by the Optimal Classification method in two dimensions, and the estimates obtained by Martin and Quinn. First I briefly explain each of the methods.

Mathematically, the Singular Value Decomposition of a rectangular matrix $\mathbf{X}_{419*9}$ is

$$\mathbf{X}_{419*9} = \mathbf{U}_{419*419}\mathbf{D}_{419*9}\mathbf{V}_{9*9} \text{ s.t. } \mathbf{U}^t\mathbf{U} = \mathbf{I} \text{ and } \mathbf{V}^t\mathbf{V} = \mathbf{I}.$$

The matrix $\mathbf{X}$ contains the original data of zeroes (dissents) and ones (votes with the majority), each case in a row and each justice in a column. This original data is decomposed into two orthogonal matrixes and a diagonal matrix. The vectors in the square matrix $\mathbf{V}$ represent the estimates of the ideal point of each justice in nine new dimensions, such that the estimates for the first dimension represents the best fit to the original data with only one dimension; the estimates for the second dimension are the best fit adding a second dimension but taking the estimates for the first dimension as given, and the estimates for the k-th dimension are the best fit in k dimensions taking the previous $k-1$ dimensions as given. Here, "best fit" means the approximation that minimizes the sum of the squared error between the approximation and the original data. The "single values" along the diagonal of $\mathbf{D}$ are all positive and represent the weights of each of the dimensions. See Eckart and Young [13] for the original mathematical idea.

The Single Value Decomposition generates nine new coordinates capturing the most frequent alignments of voting in the Court, and gives the location of each justice in all nine dimensions, so that taking only the first one or two dimensions gives the best approximation of the location of the justices in this reduced subspace.

The Eigen Decomposition and the Optimal Classification method require some previous steps. First, I calculate the disagreement matrix, which is a 9 by 9 matrix that shows for each pair of justices, the proportion of cases in which they do not vote together. Second, I convert the disagreement score matrix into a matrix of squared distances, just by squaring each cell. Third, I double-center the squared distances matrix by subtracting from each cell the row mean and the column mean, adding the matrix mean, and dividing by (-2). Double centering the squared distances matrix removes the squared terms and produces a cross-product matrix of the legislator coordinates. For details of these steps, see Poole [26]. The Eigen Decomposition of the cross-products matrix produces nine eigenvectors, which we can interpret as estimates

[7] See Brazill and Grofman [8] for a comparison of the relative merits of mutidimensional scaling methods versus factor analysis methods such as Eigen or Single Value Decomposition.

of the location of the justices in nine dimensions, and nine corresponding eigenvalues, which assign weights to each of the dimensions. Mathematically, the Eigen Decomposition of a square matrix $\mathbf{X}_{9*9}$ is

$$\mathbf{X}_{9*9} = \mathbf{U}_{9*9}\ \mathbf{D}_{9*9}\ \mathbf{U}_{9*9}^{-1},$$

where the elements of the diagonal are the eigenvalues, and the vectors of $\mathbf{U}$ the eigenvectors.

The Optimal Classification method in one dimension applied to the Supreme Court data ranks justices from one to nine, and ranks each case in between a pair of justices, predicting that all justices to one side will vote one way, and all justices on the other side will vote the other way. For instance, if a case is ranked between 2 and 3, the OC method predicts that justices 1 and 2 vote in the minority and the other seven justices in the majority. If in the real data justice 3 also votes with 1 and 2, then that's one classification error and the OC method aims to minimize the number of these errors.

The algorithm used in the Optimal Classification method is as follows. Starting with the rank ordering of the justices given by the first vector of the Eigen Decomposition of the double-centered squared-distances matrix, assign a rank to every case in such a way that the ranks minimize the total number of errors. Then, given the rank of every case, assign a new rank to the justices to minimize the number of errors given the ranking of cases. The algorithm proceeds iteratively re-ranking cases given the ranking of justices and then re-ranking justices given the ranking of cases until it converges to a solution that jointly gives a rank of both justices and cases that minimizes the number of classification errors. In two dimensions, instead of rank orderings, the OC method assigns a position in the space for each justice -or, more precisely, an area where the justice is located- and for each case it gives a cutting line partitioning the space into the area where it predicts that justices vote with the majority and the area where it predicts that justices vote with the minority. Poole [26] provides a careful explanation of this method.

To my best knowledge, the most complete analysis of the location of the ideal policies of recent Supreme Court justices is the Supreme Court Ideal Point Research conducted by Martin and Quinn [24] and [15], who use a probabilistic voting model and Bayesian inference to estimate the ideal policies of the justices in a unidimensional space. A particularly useful feature of their project is that they study the dynamics of the Court, and they update they results year by year at the homesite of the project at adw.wustl.edu/supct.php. I take the average of the estimates they report for the years 1995-2004. Estimates by Singular Value Decomposition and the Optimal Classification method range from minus one (most liberal) to plus one (most conservative). Martin and Quinn's estimates could take any value in the real line, but since the scaling of their estimates is arbitrary, I re-scale their estimates dividing by five to ease the comparison across rows in Table 4.

As shown in the table, the different methods produce similar estimates that mostly corroborate the initial ordering of the justices according to the proportion of liberal votes cast, as coded by Spaeth [34].

	1.Ste	2.Gin	Sou	Bre	5.O'Co	6.Ken	7.Reh	Sca	Tho
SVD	−0.425	−0.382	−0.351	−0.335	0.089	0.154	0.294	0.398	0.459
Eigen D	−0.418	−0.296	−0.250	−0.253	0.161	0.212	0.348	0.455	0.459
OCM 1D	$1st$	$2nd$	$4th$	$3rd$	$5th$	$6th$	$7th$	$8th$	$9th$
OCM 2D	−0.736	−0.583	−0.506	−0.498	0.169	0.274	0.489	0.704	0.661
M&Q	−0.590	−0.302	−0.248	−0.221	0.099	0.146	0.289	0.598	0.678

Table 4: Estimates of the location of the ideal policies of the justices.

The ordering according to Martin and Quinn and according to the Single Value Decomposition (SVD) coincides exactly with the ordering according to the proportion of liberal votes. It is important to note that the estimates from the SVD in the table correspond to the second dimension of the SVD. The first dimension is an "agreement dimension" in which all justices get a very similar value; this dimension captures the insight that justices tend to vote together very frequently and it is only the second vector that provides the relevant information of the location of the justices in the dimension of interest. I report the estimates for the first dimension and the weight for all nine dimensions in the appendix. Sirovich [32] used the same method to study the voting patterns of the Court from 1995 to 2002, and his estimates are similar to mine as was to be expected, with two differences. First, he fails to omit the unanimous decisions. As a consequence, the first dimension in his analysis is more accurately an agreement dimension in which all justices get an approximately equal estimate, and this (uninteresting) "agreement dimension" carries more weight than in my analysis. Second, in my data Justice Souter appears to be more liberal. This reflects the fact that Justice Souter gradually drifted during his tenure in the Court, a fact also recorded by Martin and Quinn [24].

The estimates according to the first eigenvector of the Eigen Decomposition of the cross-product of justices' coordinates switch the positions of Souter and Breyer by a very slim margin, and otherwise coincide with the proportion of liberal votes or the estimates by SVD.

The Optimal Classification method with one dimension again switches the ordering of Souter and Breyer, but with two dimensions, Optimal Classification returns Souter back to the left of Breyer and it alters the ordering of Scalia and Thomas.

All estimates agree in the following partial order $\prec$:

$$Ste \prec Gin \prec \frac{Sou}{Bre} \prec O'Co \prec Ken \prec Reh \prec \frac{Sca}{Tho}.$$

Only the relative ordering of Breyer and Souter, and the relative ordering of Scalia and Thomas remain in doubt. Rather than making a questionable assumption about these two pairs of justices, I consider all four lineal orders consistent with the partial order $\prec$ and I evaluate all the voting bloc structures that are connected according to one of these four lineal orders. Formally, a partial order is a binary relation that is reflexive, transitive and antisymmetric. A lineal order adds the property of being total, that is, it orders every pair of elements. For instance, the coalition $C = \{Ste, Gin, Sou\}$ is connected given the partial order $\prec$ because it is connected given the linear order that ranks Souter third and Breyer fourth and the coalition C' $= \{Ste, Gin, Bre\}$ is also connected given $\prec$ because it is connected given the linear order that ranks Breyer third and Souter fourth. But if a coalition contains both Gin and $O'Co$, then it must contain both Bre and Sou to be connected given $\prec$.

3.3 Endogenous Voting Blocs in the US Supreme Court

> "People ask me whether I was sorry that I was in the minority in Bush vs Gore. 'Of course I was sorry!' I'm always sorry when I don't have a majority." Justice Stephen Breyer of the US Supreme Court.[8]

To calculate the effect of voting blocs upon the utility of the justices, it is necessary to make an assumption about the utility function of the justices.

I assume that justices are outcome oriented: Each individual justice has policy preferences over the outcome of each decision, and, as quoted from Justice Breyer, wants the Court to reach a decision according to the preference of the justice. This assumption is consistent with the attitudinal model of the Court by Segal and Spaeth [31], who consider competing models of the functioning of the Court and conclude that a model of sincere voting by policy-oriented justices best explains the decisions of the Court. In earlier work, Rohde [29], studied the formation of coalitions in the writing of opinions in the Warren Court (1953-1968) and assumed that the optimization problem of the justices is to have the policy output of the Court approximate as closely as possible his own preference. If Segal and Spaeth [31] are correct and justices vote sincerely, then each justice wanted the decision of the Court to coincide exactly with the vote that the justice cast and every dissent is a defeat. If justices did not always vote sincerely, it would be difficult to discern the true preferences of the justices beyond their revealed preferences, so I assign utilities according to the actual votes cast by the justices.

The Court makes a binary decision on the merits of each case: It either affirms the ruling from a lower court, or it reverses it; it sides with the plaintiff, or with the defendant; with the liberal position, or the conservative one. For instance, in a case in which a lower court took a conservative view and sided with the plaintiff, the outcome of the decision is either affirm-plaintiff-conservative or reverse-defendant-liberal. I assume that each justice prefers one of these two outcomes over the other and each justice gets a higher utility if his preferred outcome is the one selected by the Court by majority voting. Then I assume that for the aggregate of all 419 cases from 1995 to 2004 the goal of each justice was to maximize the number of cases in which the decision of the Court coincides with the preference of the justices, as revealed by the vote of the justice. Table 3 then provides the ultimate satisfaction of each justice with the series of decisions of the Court: 419 minus the number of dissents is my measure of the utility or satisfaction of each justice with the output of the Court from 1995 to 2004. This measure of utility implicitly assumes that justices only care about how often they obtain a majority, or in other words, that they do not care more about some decisions over others. While this

[8] From "Breyer's Big Idea", by Jeffrey Toobin, in *The New Yorker*, October 31st, 2005, pages 36-43.

assumption is admittedly unrealistic, it is a simplifying step to circumvent the need to assign weights for each case and justice.

I calculate how the outcomes would have changed if justices had formed voting blocs, and how the satisfaction of each justice would have changed accordingly. For any given configuration into voting blocs in the Court, I assume that each bloc holds a private internal vote before the division of the Court, and in these internal votes I assume that each justice votes according to how the justice voted in reality in that case. Then I aggregate the votes inside each bloc according to the majority rule of the bloc, and I calculate the new outcome in the division of the Court, once I take into account that some justices now cast a vote against their preference along the lines dictated by the majority of their bloc. Finally, I calculate how many decisions change with the voting blocs under consideration relative to the original data, and for each justice I calculate the net balance of decisions that change to favor her preferences minus the number of decisions that change against her preference.

This application fits into the model detailed in the previous section, taking the frequency distribution of roll call votes in the data as the probability distribution over preference profiles in each of the cases studied by the Court. That is, if a given division of the Court occurred 10 times out of 419 cases, I assume that in each of the 419 stages of the game, this particular division of preferences occurs with probability 10/419. Since I assume that the probability distribution over preferences is invariant over time, I can use Theorem 3 to conclude that the voting blocs that emerge in the equilibrium of the one-stage game, are also the outcome of stationary equilibrium strategies in the repeated game. Therefore, for the remainder of the paper I treat the data as if justices made a single decision, once, about the formation of voting blocs, and then iterated these strategies for the duration of the repeated game.

Example 3 *Suppose Ginsburg, Souter and Breyer form a voting bloc. Then the net change in the number of decisions in which each justice is satisfied with the outcome is as follows:*

Bloc	*1.Ste*	*2.Gin*	*3.Sou*	*4.Bre*	*5.O'Co*	*6.Ken*	*7.Reh*	*8.Sca*	*9.Tho*
{234}	*12*	*12*	*4*	*4*	*-2*	*-14*	*-10*	*-12*	*-14*

Example 3 shows that had Ginsburg, Souter and Breyer committed to always vote together rolling internal dissent among the three, each of them would have achieved their preferred outcome more often even if sometimes they had to vote against their preference. Comparing these numbers to those in Table 3, Ginsburg would reduce the number of cases that end up against her preference by almost 8%. Souter and Breyer by about 3%.

If justices Ginsburg, Souter and Breyer had formed a voting bloc, 20 decisions out of 419 would have been reversed, *Atwater vs City of Lago Vista* (2001) among them. In a 5-4 decision, the Court held that the Fourth Amendment does not forbid a warrantless arrest for a minor criminal offense, such as a misdemeanor seatbelt violation punishable only by a fine. Justices Souter, Kennedy, Rehnquist, Scalia and Thomas voted with the majority. Justice O'Connor, joined by Stevens, Ginsburg and Breyer, wrote a dissent arguing that a seatbelt violation is not a reasonable ground for arrest, and thus the arrest is in violation of the Fourth Amendment that prohibits unreasonable seizure. With the exception of Souter, there is a clean division of the Court between more liberal justices favoring broader Civil Rights, and more conservative justices favoring Law Enforcement. Had Souter voted with Ginsburg and Breyer, the Court would have found the arrest to be unconstitutional.

More recently, in two famous cases decided on June 27, 2005, the Court ruled that the display of the Ten Commandments in two courthouses in Kentucky is in violation of the First Amendment Establishment Clause for the Separation of Church and State, but it also ruled that a display of the Ten Commandments in the Texas State Capitol is not unconstitutional. Justices Stevens, Ginsburg, Souter and O'Connor voted against the displays both in the Kentucky and Texas cases, while justices Kennedy, Rehnquist, Scalia and Thomas voted in favor of the displays in both cases. Justice Breyer voted against the Kentucky displays in *McCreary County vs ACLU*, giving the liberals a 5-4 majority, but he voted in favor of the Texas display in *Van Orden vs Perry*, giving the conservatives a 5-4 majority. Had Breyer voted with Souter and Ginsburg in both cases, the Texas display would have been ruled unconstitutional, just as the Kentucky ones.

Note that when a justice in a voting bloc has to vote against his true preference in the division of the Court, he would only be satisfied with the outcome if his vote -along with the whole bloc he belongs to- ends up in the minority of the Court. Hence Example 3 doesn't measure the extra number of times that Ginsburg, Souter or Breyer are in the majority, but the extra number of times that they are satisfied with the outcome. In so far as justices are ideologically motivated, it is reasonable to say that for a justice *to win* means that the preferred outcome of this justice prevails, regardless of whether the justice voted for or against her favored outcome in the division of the Court.

Epstein and Knight [14] argue that justices make strategic choices deviating from their preference for the sake of achieving the policy outcomes they desire so that the Law that emanates from the Supreme Court rulings is "the long term product of short-term strategic decision-making." I argue that if justices are strategic in their actions, then they must be tempted to form voting blocs. For instance, if Justice Breyer had formed a voting bloc with Ginsburg and Souter and no other justice had reacted to that bloc, Justice Breyer would have lost fewer cases, exactly four less.

Assume the counterfactual that Ginsburg, Souter and Breyer form a voting bloc. All three members benefit from joining so none would want to deviate and leave, disbanding the bloc. However, the formation of a single voting bloc by Ginsburg, Souter and Breyer form is not an equilibrium strategy, because other justices have incentives to react to this bloc. Table 5 displays the net payoffs to each justice relative to the benchmark with no voting blocs if Stevens joins the bloc (first row), and if the Rehnquist, Scalia and Thomas form another voting bloc (second row). A summary comparison between Table 5 and the table in Example 3 reveals that Stevens would benefit if he joined the liberal bloc, increasing his net utility from +12 to +13. Hence the voting bloc $Gin - Sou - Bre$ is not a Nash equilibrium. The second row reveals that Rehnquist, Scalia and Thomas would reduce their loses from the formation of the

$Gin - Sou - Bre$ bloc if they formed their own bloc.

Blocs	1.Ste	2.Gin	3.Sou	4.Bre	5.O'Co	6.Ken	7.Reh	8.Sca	9.Tho
{1234}	13	11	9	5	1	−15	−17	−13	−13
{234},{789}	5	−5	−7	−3	9	−1	−7	−3	−9
{123},{789}	−1	−5	−5	3	9	−3	−7	−1	−11
{789}	−5	−9	−13	−9	1	1	1	13	3
{123}	8	6	2	8	0	−10	−12	−6	−12

Table 5: Net change in satisfaction

A single bloc with the four liberal justices can occur in a Nash equilibrium, but this equilibrium does not satisfy the split-proof refinement, because Rehnquist, Scalia and Thomas have an incentive to form their own bloc to counterbalance the four liberals just as much as they do against a bloc of three liberals.

On the other hand, the formation of both a liberal bloc with Ginsburg, Souter, Breyer and a conservative bloc with Rehnquist, Scalia, Thomas is a split-proof equilibrium outcome. With the partial ordering $\prec$ discussed above, the following result summarizes my findings on hypothetical voting blocs in a nine agent assembly whose members faced an agenda and preference profile identical to those of the US Supreme Court justices between 1995-2004.

Result 4 *There exist split-proof equilibria such that any three of the four most liberal justices (Stevens, Ginsburg, Souter, Breyer) form a voting bloc and Rehnquist, Scalia and Thomas form a second voting bloc. There is no other split-proof equilibrium with connected voting blocs.*

I show the payoffs -net changes in the number of cases that each justice wins- given that $\{Ste, Gin, Sou\}$ form a voting bloc, and the three most conservative justices form another voting bloc in row three of Table 5. For the sake of comparison, in rows four and five I show the payoffs if one member of the liberal bloc deviates and the bloc dissolves, leaving $\{Reh, Sca, Tho\}$ as the unique bloc, and the payoffs if the conservative bloc dissolves and $\{Ste, Gin, Sou\}$ remain as a bloc. It is clear that no justice wants to abandon the bloc he belongs to.

In the appendix I provide a table with the payoffs for each justice for a sample of voting blocs, and for the two configurations of the assembly into connected voting blocs that are supported in a split-proof equilibrium not already listed in Table 5. For all other configurations with connected voting blocs, I provide the payoffs for each justice and a deviation (if any exists) that makes the corresponding strategy not a best response in an Excel file that also contains the original data and formulas to replicate the calculations. This file is available from the author.

According to Result 4, the equilibrium outcomes are such that two opposing blocs -one at each side of the ideological spectrum- counterbalance each other, and the swing moderate agents, in this case O'Connor and Kennedy remain unaffiliated, independent. Voting blocs merely reinforce the polarization of the Court into a liberal group and a conservative group of justices, and do not produce a major realignment of votes. As a curiosity, the most famous decision of this court, the 5-4 division in Bush vs Gore (2000) which stopped the recount of the Florida votes and gave Bush the presidency would not have been reversed, since it was already the case that the four liberals voted together in the minority, and the two moderate conservatives and three conservatives voted together in the majority. The formation of permanent connected voting blocs would have made this particular 5-4 conservative-liberal division more frequent, but this was already the most frequent division of the Court.

Consider the voting bloc configuration in which Stevens, Ginsburg and Souter form a voting bloc, and Rehnquist, Scalia and Thomas form another voting bloc. In terms of the location of each of the blocs in an ideological space, each of the blocs converges near the location of its median member. Indeed, the liberal bloc casts 313 liberal votes where Ginsburg alone casts 308 (Stevens and Souter cast 344 and 307, see Table 3) and its position in space according to SVD is -0.318, indistinguishable from Ginsburg's in the dimension of interest, while the conservative bloc casts 72 liberal votes for 84 of Scalia (98 and 71 by Rehnquist and Thomas) and locates by SVD at 0.392, where Scalia alone is at 0.398.

Compare Result 4 with the stylized assembly with 9 agents who have symmetric and independent types. Note that the equilibrium predictions with the US Supreme Court data are a subset of the predictions for the idealized assembly. The theoretical model predicted two voting blocs of size three, one with three of the four most liberal members, the other with three of the four most conservative members. The prediction with the empirical data fits within this set of outcomes, and the only difference is that the conservative bloc has to be $\{789\}$ and cannot be $\{678\}$ instead. The cause of this difference is that the modelled assembly assumed that agents 3 and 4 and agents 6 and 7 are identical. In the empirical application, Souter and Breyer are indeed similar enough in their voting behavior, but Kennedy is markedly different from Rehnquist, and in particular Kennedy is not conservative enough to benefit from forming a bloc with Rehnquist and Scalia.

The following two comments suggest that Result 4 should be interpreted with caution.

First, the predicted voting blocs are in accordance to those predicted by the more abstract model and reinforce the intuition that the assembly is likely to divide into two opposing voting blocs that counterbalance each other, one at each side of the ideological spectrum and leaving a number of unaffiliated moderate independents. However, this result is based on the particular equilibrium concept that I have chosen. The question of which equilibrium refinement or which stability concept is appropriate is still open in the literature.

Second, this section has shown what voting blocs would occur in equilibrium in an assembly with nine rational agents who are strategic and can coordinate their votes without constraints, and whose preferences are consistent with those revealed by the pattern of votes in the US Supreme Court for 1995-2004. It has not provided, nor did it intend to provide, a theory of voting in the Court. I leave to Supreme Court scholars these tasks. Restraints of a legal, normative or ethical nature may deter Supreme Court justices from committing to vote as a bloc and this section doesn't attempt to explain voting in the US Supreme Court as much as it intends to illustrate how voting blocs affect outcomes in practice, and what voting blocs we expect to arise in an assembly of political agents who coalesce strategically. The data from the US Supreme Court serves by proxy to shed some light into the formation of voting blocs in committees, councils, small assemblies, and all sorts of political caucuses, in which the incentives to form the blocs will be salient and the restraints that Supreme Court justices face are probably absent -and, crucially, the data on the preferences of its members is also absent.

I have proved that members of a committee or assembly with size and preferences identical to those of the US Supreme Court face strategic incentives to coalesce into voting blocs. An explanation of whether or not the US Supreme Court justices act upon these strategic incentives is beyond the scope of this paper.

4 Conclusion and Extensions

Members of a democratic assembly -legislature, council, committee- can affect the policy outcome by forming voting blocs. A voting bloc coordinates the voting behavior of its members according to an internal voting rule independent of the rule of the assembly, and this coordination of votes affects the outcome in the division of the assembly.

I have shown that for any assembly there exists an equilibrium in which agents endogenously form voting blocs and these blocs affect voting records exercising party discipline. I have further shown that permanent voting blocs -parties- persist over time in equilibrium until there is a change in the underlying heterogeneous preferences.

In a small assembly, I predict that two polarized voting blocs form. I first derive this prediction from an abstract assembly with stylized preferences, but I obtain the same prediction when I assign preferences according to the empirical data on voting records in the United Supreme Court between 1995 and 2004. In either case, the outcome of a coalition formation game in which agents strategically formed voting blocs would result in two voting blocs, one at each side of the median member of the assembly.

I have shown that strategic members of an assembly form political parties to coordinate their votes. I explain political parties as endogenous voting blocs in an assembly. This motivation to form parties is complementary to the traditional views of parties as electoral machines, but it is simpler, because it does not rely on outside agents like citizens or lobbies, and it is also more general, since it can also explain the formation of parties in unelected assemblies, such as the House of Lords in the United Kingdom, the formation of blocs in the assemblies of international organizations such as the UN or the IMF, or the formation of European parties in the European Parliament, parties that vote together but nevertheless fracture into independent national parties for electoral purposes.

The theory in this paper has several natural extensions: Comparing the results under other equilibrium concepts, such as Coalition-Proofness or Equilibrium Binding Agreements; allowing for a richer class of rules, not just anonymous and majoritarian rules; introducing intensity of preferences so that agents who like a proposal do so to varying degrees; considering unequally weighted individuals or even pyramidal structures, in which individual agents coalesce into factions, factions coalesce into parties (voting blocs of second order), parties into alliances (voting blocs of third order) and so on... Empirical applications range from revisiting the historical records of the early United States Congress to try to determine the incentives to coordinate votes along State lines or along parties, to salient current developments such as the theoretical advantages to each of the 27 European Union countries from pooling their votes under a common foreign EU policy. These questions constitute an agenda for further research.

5 Appendix

5.1 Estimates by Eigen-D, SVD and OCM-2D

Single values	1.Ste	2.Gin	Sou	Bre	5.O'Co	6.Ken	7.Reh	Sca	Tho
1st-Eigenvector	−0.418	−0.298	−0.250	−0.253	0.161	0.212	0.348	0.455	0.459
2nd-Eigenvector	0.339	0.026	0.198	−0.467	−0.493	−0.146	−0.348	0.277	0.406
SVD 1st Dim	−0.241	−0.306	−0.331	−0.326	−0.404	−0.397	−0.359	−0.301	−0.304
SVD 2nd Dim	−0.425	−0.382	−0.351	−0.335	0.089	0.154	0.294	0.398	0.402
OCM 1st Dim	−0.736	−0.583	−0.506	−0.498	0.169	0.274	0.489	0.704	0.661
OCM 2nd Dim	0.569	0.087	0.387	−0.730	−0.710	−0.182	−0.524	0.482	0.661

Dimension i	$1st$	$2nd$	$3rd$	$4th$	$5th$	$6th$	$7th$	$8th$	$9th$
Eigenvalue α_i	−0.817	0.082	0.044	0.037	0.022	−0.021	0.015	−0.008	0.006
Weight Dim λ_i SVD	0.385	0.179	0.079	0.072	0.067	0.062	0.056	0.052	0.047

The top table contains the first and second eigenvectors obtained by the Eigen Decomposition of the double-centered matrix of squared distances of the justices, the estimates of the location of the justices in the first and second dimension by Single Value Decomposition and the estimates of the location in the first and second dimensions obtained by the Optimal Classification method with two dimensions. Note that the first dimension with SVD is an "agreement dimension" where all justices take a similar position, and it is only the second dimension that is the relevant and meaningful one, comparable to the first dimension in the other methods. The bottom table provides the nine single values of the Eigen Decomposition, and the weights of the nine dimensions from the SVD (to obtain the single value of each dimension, multiply by 11.065).

5.2 Table following Result 4

Blocs	1.Ste	2.Gin	3.Sou	4.Bre	5.O'Co	6.Ken	7.Reh	8.Sca	9.Tho
$\{124\}, \{789\}$	−1	−5	−1	−1	9	−3	−7	−1	−7
$\{134\}, \{789\}$	−1	1	−7	−3	11	−3	−9	−1	−5
$\{345\}$	75	77	71	85	−67	−79	−77	−85	−89
$\{456\}$	−4	−4	−10	6	14	2	−2	−8	−16
$\{567\}$	−50	−48	−52	−44	−12	16	48	46	44
$\{4567\}$	−8	−4	−10	−2	6	6	10	0	−8
$\{34567\}$	−15	−15	−11	7	17	9	11	−15	−23

The first column in each row contains the voting bloc structure as a list of the blocs that form; the numbers inside each bloc correspond to the justices in the order given in the top row. The other cells detail the payoff to each justice. The first two rows correspond to the voting blocs that appear in two different split-proof equilibrium outcomes. The others are not part of any split-proof equilibrium outcome.

References

[1] John H Aldrich. *Why Parties? The Origin and Transformation of Political Parties in America*. Chicago, University of Chicago Press, 1995.

[2] Robert M Axelrod. *Conflict of Interest: A Theory of Divergent Goals with Application to Politics*. Chicago, Markham, 1970.

[3] David Baron. Government formation and endogenous parties. *American Political Science Review*, 85:137–164, 1993.

[4] David Baron and Ehud Kalai. The simplest equilibrium of a majority rule division game. *Journal of Economic Theory*, 61:290–301, 1993.

[5] B. Douglas Bernheim, Bezalel Peleg, and Michael D Whinston. Coalition-proof nash equilibria: I. concepts. *Journal of Economic Theory*, 42(1):1–12, 1987.

[6] Francis Bloch. *The Endogenous Formation of Economic Coalitions*, chapter Non-Cooperative Models of Coalition Formation in Games with Spillovers. Edward Elgar Publishing, 2003.

[7] Francis Bloch and Armando Gomes. Contracting with externalities an outside options. *Journal of Economic Theory*, 127(1):171–201, 2006.

[8] Timothy J. Brazill and Bernard Grofman. Factor analysis versus multi-dimensional scaling: Binary choice roll-call voting and the us supreme court. *Social Networks*, 24:201–229, 2002.

[9] Carlo Carraro. *The Endogenous Formation of Economic Coalitions*. Edward Elgar Publishing, 2003.

[10] Gary W Cox and Matthew D McCubbins. *Legislative Leviathan*. Berkeley, University of California Press, 1993.

[11] Daniel Diermeier and Antonio Merlo. Government turnover in parliamentary democracies. *Journal of Economic Theory*, 94(1):46–79, 2000.

[12] John Duggan and Mark Fey. Repeated downsian electoral competition. *International Journal of Game Theory*, 35(1):39–69, dec 2006.

[13] Carl Eckart and Gale Young. The approximation of one matrix by another of lower rank. *Psychometrika*, 1:211–218, 1936.

[14] Lee Epstein and Jack D Knight. *The Choices Justices Make*. Washington DC, CQ Press, 1998.

[15] Lee Epstein, Andrew D. Martin, and Kevin M. Quinn. The rehnquist court? *Law and Courts*, 15:18–23, 2005.

[16] Dan S Felsenthal and Moshé Machover. Annexations and alliances: When are blocs advantageous a priori. *Social Choice and Welfare*, 19:295–312, 2002.

[17] Drew Fudenberg and Jean Tirole. *Game Theory*. Cambridge, MIT Press, 1991.

[18] Andrew Gelman. Forming voting blocs and coalitions as prisoner's dilemma: A possible theoretical explanation for political instability. *Contributions to Economic Analysis and Policy*, 2(1), 2003.

[19] Kyle Hyndman and Debraj Ray. Coalition formation with binding agreements. *Review of Economic Studies*, forthcoming.

[20] Matthew Jackson and Boaz Moselle. Coalition and party formation in a legislative voting game. *Journal of Economic Theory*, 103(1):49–87, 2002.

[21] Marek M. Kaminski. Coalitional stability of multi-party systems: Evidence from poland. *American Journal of Political Science*, 45:294–312, 2001.

[22] Marek M. Kaminski. A general equilibrium model of multi-party competition. *Social Choice and Welfare*, 26:333–361, 2006.

[23] Gilat Levy. A model of political parties. *Journal of Economic Theory*, 115(2):250–277, 2004.

[24] Andrew D. Martin and Kevin M. Quinn. Dynamic ideal point estimation via markov chain monte carlo for the u.s. supreme court, 1953-1999. *Political Analysis*, 10(2):134–153, 2002.

[25] Massimo Morelli. Party formation and policy outcomes under different electoral systems. *The Review of Economic Studies*, 71(3):829–853, 2004.

[26] Keith Poole. *Spatial Models of Parliamentary Voting*. Cambridge University Press, 2005.

[27] Debraj Ray and Rajiv Vohra. Equilibrium binding agreements. *The Journal of Economic Theory*, 73:30–78, 1997.

[28] William H Riker. *The Theory of Political Coalitions*. Yale University Press, 1962.

[29] David W Rohde. Policy goals and opinion coalitions in the supreme court. *The Midwest Journal of Political Science*, 16(2):208–224, 1972.

[30] Norman Schofield and Itai Sened. *Multiparty Democracy*. Cambridge University Press, 2006.

[31] Jeffrey A Segal and Harold J Spaeth. *The Supreme Court and the Attitudinal Model Revisited*. Cambridge: Cambridge University Press, 2002.

[32] Lawrence Sirovich. A pattern analysis of the second rehnquist u.s. supreme court. *Proceedings of the National Academy of Sciences*, 100(13):7432–7437, 2003.

[33] James Snyder and Michael Ting. An informational rationale for political parties. *American Journal of Political Science*, 46:90–110, 2002.

[34] Harold Spaeth. United states supreme court judicial database 1953-2005. Computer File available at www.as.uky.edu/polisci/ulmerproject, apr 2006.

[35] Robert M. Thrall and William F Lucas. N-person games in partition form. *Naval Research Logistics Quarterly*, 10:281–298, 1963.

[36] Sang-Seung Yi. *The Endogenous Formation of Economic Coalitions*, chapter Endogenous Formation of Economic Coalitions: A Survey of the Partition Function Approach. Edward Elgar Publishing, 2003.