Political compromise and endogenous formation of coalitions[*]

Anke Gerber[1], Ignacio Ortuno-Ortin[2]

[1] IMW, University of Bielefeld, Postfach 10 01 31, D 33501 Bielefeld, Germany
[2] Department of Economics and IVIE, University of Alicante, E 03080 Alicante, Spain

Abstract. A one-dimensional model of spatial political competition with endogenous party formation is developed. It is proved that at equilibrium there are only two parties. These parties propose alternatives in the extreme positions of the policy space. The adopted policy, however, is a compromise between these two extremes.

1 Introduction

The spatial framework first suggested by Hotelling (1929) to model political competition has provided a set of important results and insights. The best known of these results is the ''principle of minimum differentiation.'' During the last two decades many modifications of the original Hotelling model have been developed. The basic insight, however, is still widely accepted: parties have strong incentives to converge in their proposed policies. The best proof of this is the extended use of the median voter theorem in many recent papers on Political Economy.

An essential assumption in this literature is that the party that obtains more than half of the total number of votes has the power to implement its

announced policy. Recently, some authors have challenged this assumption. Alesina and Rosenthal [1] develop a model of political spatial competition with the distinctive feature that the implemented policy is the result of a compromise between the President and the Congress. Even though most of their work concentrates on the US system, they also claim that the outcome of the political competition in parliamentary countries can generally be seen as a compromise between the different political forces. One of the main conclusions here is that in equilibrium parties' policies do not converge. Moreover, it may be the case that parties adopt very radical policies. The authors also provide empirical evidence of party polarization in the USA. Ortuno-Ortin [7] studies the existence of an equilibrium in a model in which the outcome of the game also depends on the distribution of votes among the two parties.

In the present paper we also take the view that democratic institutions achieve a "compromise" between all the political proposals. We capture this idea by assuming a spatial model in which the implemented policy is a weighted convex combination of all the proposed policies. The weights will depend on the distribution of votes among the proposals. Different democratic institutions and societies will translate votes into weights in different ways. One example would be Proportional Representation under which each proposal gets a weight proportional to the share of votes that it received. Our analysis, however, allows for very general ways to calculate such weights. Thus, we also consider "weight functions" in which, for example, the proposal that gets more than half of the votes obtains a very large weight and proposals with a small share of the votes get an arbitrarily small weight.

We do not provide a theory about the way votes are translated into weights. But we will be general enough to be able to capture many possible scenarios. For example, we could see the implemented policy as the result of some bargaining process among different parties, each of them defending its announced policy. The bargaining power of each party may depend on the distribution of votes. In this case the implemented policy can be seen as a combination of the different proposals.

We will impose, however, two main restrictions on the allowed weight functions: continuity and superadditivity. The first one implies that the weights associated with the proposals change in a continuous way with respect to the distribution of votes. Thus, institutions in which "the winner takes all" are not considered here. The second restriction, superadditivity, is weaker. Say that one proposal gets weight w_1 and a second gets w_2. Then, if all the agents that support these two proposals would agree to vote for only one proposal, the weight of this proposal would be at least $w_1 + w_2$.

We take the "compromise" approach to political competition to investigate the endogenous formation of parties or coalitions. Any element of the feasible policy space (which is exogenously given) can be a political proposal. Parties can be seen as coalitions of agents putting forward specific proposals. At equilibrium there is no coalition of agents such that if they

change strategies and vote for different proposals none of them is worse off and at least one of them is better off. Thus, our equilibrium concept will be strong Nash.

The paper shows that if the policy space is a one-dimensional interval there exists a unique strong Nash equilibrium. The equilibrium is such that the population is divided into two sets and all agents in one of them vote for the smallest policy in the interval and all the other agents vote for the greatest policy in the interval. The interpretation is that in the one-dimensional case there is only "room" for two parties. Furthermore, parties are polarized (but the implemented policy is not). The intuition is clear; the implemented policy must be a compromise among the different proposals (convex combination of the policies that obtained some votes). If there were, say, three proposals the "compromise" policy could be moved in the desired direction by two of the coalitions joining forces together. Moreover, voting for a radical policy also moves the compromise in that direction.

This extreme polarization of the equilibrium proposals may be criticized as very unrealistic. Notice, however, that the policy space was exogenously given and it might be the case that the end points of such space are not, in reality, too radical. For example, imagine that the policy space is the interval [0, 1] which represents possible income taxation policies. Then the point 0 could be understood as the lowest possible taxation level that can realistically be proposed (this normally changes with time), say 10%. The point 1 could be seen, for example, as an average taxation level of 60%

In any case, our results emphasize the incentives to make radical proposals and this is in contrast to the standard Hotelling approach which isolates the incentives to become moderate. Reality is, no doubt, much more complex and the incorporation of elements such as asymmetric information, costs of party or coalition formation, interest groups, and so on would introduce new forces and incentives which could give a more realistic picture of political competition.

Fedderson [3] also considers a model of party formation. Here, however, the policy that receives most votes wins (and is implemented). Voting is costly and each agent may stay home or vote for any policy. The number of agents is finite so that each of them might be pivotal. It is shown that in any Nash equilibrium at most two policies obtain votes. Thus, even though this result resembles the one in our model, the differences are important. No agent has "positive mass" in our model and we show existence of a unique strong Nash equilibrium. Furthermore, the outcome of the game does not coincide with the proposal of the winner. The intuition behind the results in Feddersen is, consequently, completely different from the ones in our model. It should also be mentioned that the results in Greenberg and Weber [4], [5], and Le Breton and Weber [2] do not apply to our model. These authors consider some problems of stable coalition structures and the equilibrium concepts are similar to the ones in this paper. The models are, however, quite different. In their case each possible coalition of agents has a set of achievable alternatives. They can implement one of those alternatives regardless of what

other coalitions do. In our case *one* alternative is implemented and coalitions smaller than the grand coalition only have partial influence in choosing this alternative.

The organization of the rest of the paper is as follows. Section 2 sets up the basic model. The main results of the paper are presented in Section 3. Section 4 concludes and discusses some modifications of the model.

2 The voting game

Let $\mathscr{B}$ denote the set of Borel sets in $[0, 1]$ and let λ denote the Lebesgue measure on $\mathscr{B}$. For $A \subset [0, 1]$ let A^c denote the complement of A in $[0, 1]$. Let $\mathbb{R}_+ = \{x \in \mathbb{R} | x \geq 0\}$.

We assume that the set of agents' type is the interval $[0, 1]$. A set $A \in \mathscr{B}$ is called a *coalition* of types. The distribution of types is described by a measurable function $f : [0, 1] \rightarrow \mathbb{R}_+$, such that $\int_0^1 f \, d\lambda = 1$. The policy space is given by the set $[0, 1] \cup \{-1\}$. The element $\{-1\}$ in the policy space can be seen as representing a very bad outcome for cases of "total disagreement." It is only needed for technical reasons since we want to exclude those cases in which agents form non-measurable coalitions.[1]

Agents have preferences over the policy space. The preferences of an agent of type x are represented by $u_x : [0, 1] \cup \{-1\} \rightarrow \mathbb{R}_+ \cup \{-1\}$, where $u_x(-1) = -1$, u_x is single-peaked, and $u_x(x) > u_x(y) \geq 0 \ \forall y \in [0, 1], y \neq x$. Thus, x is the ideal policy for an agent of type x. Society will have to choose a unique policy from the policy space. In order to do that each agent will announce an element of this space. Thus, $S_x = [0, 1]$ is the *strategy space* of type x, $x \in [0, 1]$, and we interpret $s_x \in S_x$ as type x voting for policy s_x. We define $S = \prod_{x \in [0, 1]} S_x$.

Before defining the game form used for choosing a policy we need some additional definitions. Let v be the measure on $\mathscr{B}$ that has density f with respect to λ :

$$v(A) = \int_A f \, d\lambda \quad (A \in \mathscr{B}).$$

Thus, $v(A)$ can be understood as the "size" of coalition A. Given $s \in S$ and $c \in [0, 1]$ the coalition of types voting for c is given by

$$A_c^s = \{x \in [0, 1] \mid s_x = c\}.$$

Definition 2.1 $s \in S$ is called **admissible** if $A_c^s \in \mathscr{B} \quad \forall \, c \in [0, 1]$.

---

[1] Readers who are not interested in these technical details can think of the policy space as being the interval $[0, 1]$.

Given the admissible strategy profile s the set of *relevant proposals* C_s is defined by

$$C_s = \{c \in [0, 1] \mid v(A_c^s) > 0\} .$$

Observe that C_s is countable. Let

$$W = \{(w_n)_n \mid 0 \le w_n \le 1 \ (n \in \mathbb{N}), \sum_n w_n = 1\} .$$

If $s \in S$ is admissible then clearly $(v(A_c^s))_{c \in C_s} \in W$ [2]. The influence or weight each candidate will have in the determination of the adopted policy depends on the size of the coalition voting for it. The weights, however, do not need to be proportional to the size of the coalitions. For this purpose we define a weight function $g : W \to W$. We write $g(w) = (g_n(w))_n \quad (w \in W)$.

We impose the following assumptions on the weight function g.

A.1 (Anonymity)

Let $w = (w_n)_n \in W$ and let $\pi : \mathbb{N} \to \mathbb{N}$ be one-to-one and onto. Let $w' \in W$ be defined by $w'_{\pi(n)} = w_n (n \in \mathbb{N})$. Then

$$g_{\pi(n)}(w') = g_n(w) \quad (n \in \mathbb{N}).$$

A.2 (Zero-Property)

Let $w = (w_n)_n \in W$. Then

$$g_1(w) = 0 \Leftrightarrow w_1 = 0.$$

A.3 (Superadditivity)

Let $w = (w_n)_n \in W$. Define $w' \in W$ by $w'_1 = w_1$, $w'_2 = \sum_{n=2}^{\infty} w_n$, $w'_n = 0$ $(n \ge 3)$. Then

$$g_2(w') \ge \sum_{n=2}^{\infty} g_n(w).$$

Before stating assumption (A.4) we define a metric d on W by

$$d(w, w') = \sup_n |w_n - w'_n| \quad (w, w' \in W). [3]$$

For $(w^m)_m \subset W$ and $w \in W$ we write $w^m \to w$ if $(w^m)_m$ converges to w with respect to the metric d.

A.4 (Continuity)

Let $(w^m)_m \subset W$ and $w \in W$. If $w^m \to w$ then $g(w^m) \to g(w)$.

(A.1) is a standard anonymity assumption. (A.2) just states that candidates have an influence on the determination of the outcome if and only if they receive votes from a set of agents with positive measure. Assumption (A.3) is a form of (weak) increasing returns to the number of votes. Notice

---

[2] If $C_s = (c_1, \cdots, c_n)$ is finite we identify $(v(A_c^s))_{c \in C_s}$ with the sequence $(v(A_{c_1}^s), v(A_{c_2}^s), \ldots, v(A_{c_n}^s), 0, \ldots) \in W$.

[3] d is the standard metric in the space l_∞.

that Proportional Representation satisfies this assumption. (A.4) is the most restrictive assumption. It imposes continuity on g and, therefore, rules out games in which the "winner takes all."

In the case of only two coalitions it is natural to require that g is monotonic, i.e. the largest coalition obtains a greater weight. Even though we have not imposed monotonicity it directly follows from assumptions (A.1), (A.2) and (A.3).

Definition 2.2 *Let* $w = (w_1, w_2, 0, \ldots), w' = (w'_1, w'_2, 0, \ldots) \in W$. $g : W \to W$ *is* **monotonic** *if*

$$w_1 \leq w'_1 \Rightarrow g_1(w) \leq g_1(w').$$

Lemma 2.3 *Let* $g : W \to W$ *obey assumptions (A.1), (A.2) and (A.3). Then g is monotonic.*

Proof: Let $\epsilon \geq 0$ and $w = (w_1, w_2, 0, \ldots), w' = (w_1 + \epsilon, w_2 - \epsilon, 0, \ldots)$, $w'' = (w_1, \quad w_2 - \epsilon, \epsilon, 0, \ldots) \in W$. By assumptions (A.1) and (A.3) $g_1(w'') + g_3(w'') \leq g_1(w')$. This implies $g_1(w'') \leq g_1(w')$. Also, by assumption (A.3) $g_2(w'') + g_3(w'') \leq g_2(w)$. Therefore, by assumption (A.2) $1 - g_1(w'') \leq 1 - g_1(w)$. It follows that $g_1(w) \leq g_1(w'') \leq g_1(w')$.

<div align="right">q.e.d.</div>

Now we can specify the outcome function $F : S \to [0, 1] \cup \{-1\}$ that describes how society chooses a policy. For all $s \in S$ let

$$F(s) = \begin{cases} \sum_{c \in C_s} c \, g_c((v(A_c^s))_{c \in C_s}) & , \quad \text{if } s \text{ is admissible} \\ -1 & , \quad \text{otherwise} \end{cases} .$$

F is well defined since C_s is countable for any admissible strategy $s \in S$. Since g has the zero-property the outcome is given by the convex combination of the relevant proposals where the weights of each proposal are determined by g. An important special case would be Proportional Representation. In this case g is the identity function: the weight a candidate obtains is identical to the proportion of agents who vote for it.

Given $s \in S$ the *payoff* for type x is given by $u_x(F(s))$.

Definition 2.4 *A* **voting game** *is given by*

$$\Gamma = \left(f, \, S, \, (u_x)_{x \in [0,1]}, \, g, \, F \right) .$$

Definition 2.5 $s \in S$ *is a* **strong Nash equilibrium (SNE)** *of the voting game* Γ *if there exists no* $D \in \mathscr{B}$ *and* $s' \in S$ *such that* $s'_x = s_x \ \forall x \notin D$ *and*

$$u_x(F(s')) \geq u_x(F(s)) \quad \forall x \in D$$
$$\text{and } u_x(F(s')) > u_x(F(s)) \quad \text{for at least one } x \in D.$$

Notice that our definition of SNE allows for deviations where only a set of agents with measure zero ends up better off. This is clearly a very weak requirement and therefore the equilibrium concept is strong.

3 The main results

The following proposition is a well known result from measure theory (see for example [6], Theorem B, p. 125).

Proposition 3.1 *A finite measure v on a σ-algebra $\mathscr{B}$ is absolutely continuous with respect to a measure μ on $\mathscr{B}$ iff $\forall \epsilon > 0 \; \exists \delta > 0$ such that*

$$\mu(A) < \delta \Rightarrow v(A) < \epsilon \; (A \in \mathscr{B}) \; .$$

The following lemma shows that there exists a strategy profile $\hat{s}$ such that the outcome of the game coincides with the ideal policy of the "dividing type." Any type who is to the left (right) of $\hat{x}$ votes for 0 (1).

Lemma 3.2 *Let $\Gamma = (f, S, (u_x)_{x\in[0,1]}, g, F)$ be a voting game and let g obey the assumptions (A.1) – (A.4). Then there exists a unique $\hat{x} \in [0,1]$ such that for $\hat{s} \in S$ defined by $\hat{s}_y = 0 \, (y \le \hat{x})$, $\hat{s}_y = 1 \, (y > \hat{x})$ it is true that*

$$F(\hat{s}) = \hat{x} \; .$$

Proof: By Proposition 3.1

$$v([0,x]) = \int_{[0,x]} f d\lambda$$

is continuous as a function of x, $x \in [0,1]$. Therefore, and by the continuity of g, the function $h\colon [0,1] \to [0,1]$ defined by $h(x) := g_1(v([0,x]), 1 - v([0,x]))$ $(x \in [0,1])$ is continuous. By assumption (A.2) we have $h(0) = 1$, $h(1) = 0$. By monotonicity of g_1 the function h has a unique fixed point $\hat{x}$. Hence,

$$F(\hat{s}) = 0 \cdot g_0(v([0,\hat{x}]), 1 - v([0,\hat{x}])) + 1 \cdot g_1(v([0,\hat{x}]), 1 - v([0,\hat{x}])) = \hat{x} \; .$$

q.e.d.

We can now provide the main result of the paper. It basically states that $\hat{s}$ is the only SNE in our voting game. Thus, in the one-dimensional case there is only "room" for two coalitions (parties). Notice that, in general, the equilibrium outcome $\hat{x}$ coincides with neither the median nor the mean type.

Theorem 3.3 *Let $\Gamma = (f, S, (u_x)_{x\in[0,1]}, g, F)$ be a voting game and let g obey the conditions (A.1)–(A.4). Let $\hat{s}$ and $\hat{x}$ be the strategy and type defined in Lemma 3.2. Then $\hat{s}$ is a SNE of the voting game Γ. Moreover, $\hat{s}$ is unique in the sense that there exists no SNE $s \in S$ such that $F(s) \ne F(\hat{s})$ and if s is a SNE and $F(s) = F(\hat{s})$ then $v(A_c^s) = 0$ for all $c \in C_s$ such that $c \notin \{0, 1\}$.*

Proof. I $\hat{s}$ is a SNE for Γ.

Observe that no coalition containing types voting for 0 *and* types voting for 1 can improve by deviating from $\hat{s}$. If for example one member $x \le \hat{x} = F(\hat{s})$ of such a coalition is better off after the deviation all agents x with $x > \hat{x}$ must be worse off. Therefore, we only need to consider the case where a coalition $D \subset [0, \hat{x}]$ deviates. (The proof for $D \subset (\hat{x}, 1]$ is similar.) Let $s \in S$ be such

that $s_y = \hat{s}_y$ $(y \in D^c)$. We will show that D cannot improve by switching to s. If s is not admissible we are done since the outcome would be -1 which is worse for all $y \in D$. Therefore, let s be admissible. Then

$$F(s) = \sum_{c \in C_s} c\, g_c\left((v(A_c^s))_{c \in C_s}\right)$$

$$\geq 1 \cdot g_1\left((v(A_c^s))_{c \in C_s}\right)$$

$$= 1 - \sum_{1 \neq c \in C_s} g_c\left((v(A_c^s))_{c \in C_s}\right)$$

$$\geq 1 - g_0(1 - v(A_1^s), v(A_1^s))$$

$$= g_1\left(1 - v(A_1^s), v(A_1^s)\right)$$

$$\geq g_1\left(v([0,\hat{x}]),\ 1 - v([0,\hat{x}])\right)$$

$$= F(\hat{s})$$

by superadditivity and monotonicity of g. (Observe that $(\hat{x}, 1] \subset A_1^s$.) Therefore, $u_y(F(s)) \leq u_y(F(\hat{s}))$ $(y \in D)$ which proves that $\hat{s}$ is a SNE.

II. $\hat{s}$ is the unique SNE for Γ.
Assume first that s is a SNE such that $F(s) \neq F(\hat{s})$. If $F(s) = -1$ then switching to $\hat{s}$ would improve all types, which is a contradiction. W.l.o.g. let $-1 \neq F(s) < F(\hat{s})$. For $x \in [0,1]$ define the strategy $s^x \in S$ by $s_y^x = 1$ $(y \geq x)$, $s_y^x = s_y$ $(y < x)$. By Proposition 3.1 and by continuity of g the outcome $F(s^x)$ is continuous in x. Therefore, there are two cases.

Case (a): $\exists x \geq \hat{x}$ such that $F(s^x) = x$.
But then coalition $[x, 1]$ could improve upon s by s^x since $F(s) < F(\hat{s}) = \hat{x} \leq x = F(s^x)$. Therefore, s is no SNE Contradiction.
Case (b): $x > F(s^x)$ for all $x \geq \hat{x}$.
Then

$$F(s^{\hat{x}}) = \sum_{c \in C_{s^x}} c\, g_c\left(\left(v\left(A_c^{s^x}\right)\right)_{c \in C_{s^x}}\right)$$

$$\geq 1 \cdot g_1\left(\left(v\left(A_c^{s^x}\right)\right)_{c \in C_{s^x}}\right)$$

$$= 1 - \sum_{1 \neq c \in C_{s^x}} g_c\left(\left(v\left(A_c^s\right)\right)_{c \in C_{s^x}}\right)$$

$$\geq 1 - g_0\left(1 - v\left(A_1^{s^x}\right), v\left(A_1^{s^x}\right)\right)$$

$$\geq g_1\left(v([0,\hat{x}]),\ 1 - v([0,\hat{x}])\right)$$

$$= F(\hat{s}) = \hat{x}$$

by superadditivity and monotonicity of g which is a contradiction to $\hat{x} > F(s^{\hat{x}})$.

Assume now that $s \in S$ is a SNE with $F(s) = F(\hat{s})$ and that $\exists \bar{c} \in C_s$, $0 \neq \bar{c} \neq 1$, such that $v(A_{\bar{c}}^s) > 0$. W.l.o.g. let $v(A_{\bar{c}}^s \cap [0, \hat{x}]) > 0$. For $x \in [0, 1]$ define the strategy $s^x \in S$ by $s_y^x = 1$ $(y \geq x)$, $s_y^x = s_y$ $(y < x)$. As before the outcome $F(s^x)$ is continuous in x. By (A.2)

$$g_{\bar{c}} \left(\left(v \left(A_c^{s^x} \right) \right)_{c \in C_{s^x}} \right) > 0 \text{ for all } x \geq \hat{x} ,$$

which implies that

$$F(s^{\hat{x}}) = \sum_{c \in C_{s^x}} c \, g_c \left(\left(v \left(A_c^{s^x} \right) \right)_{c \in C_{s^x}} \right)$$

$$> 1 \cdot g_1 \left(\left(v \left(A_c^{s^x} \right) \right)_{c \in C_{s^x}} \right)$$

$$= 1 - \sum_{1 \neq c \in C_{s^x}} g_c \left(\left(v \left(A_c^{s^x} \right) \right)_{c \in C_{s^x}} \right)$$

$$\geq 1 - g_0 \left(1 - v \left(A_1^{s^x} \right), v \left(A_1^{s^x} \right) \right)$$

$$= g_1 \left(1 - v \left(A_1^{s^x} \right), v \left(A_1^{s^x} \right) \right)$$

$$\geq g_1 \left(v([0, \hat{x}]), 1 - v([0, \hat{x}]) \right)$$

$$= F(\hat{s}) = \hat{x}$$

by superadditivity and monotonicity of g. Since $F(s^x)$ is continuous in x there are again two cases.

Case (a): $\exists x > \hat{x}$ such that $F(s^x) = x$.
then $F(s^x) = x > \hat{x} = F(\hat{s}) = F(s)$ and coalition $[x, 1]$ can improve upon s by s^x. Therefore, s is no SNE which is a contradiction.
Case (b): $x > F(s^x)$ for all $x > \hat{x}$.
But then by continuity also $\hat{x} \geq F(s^{\hat{x}})$. This contradiction proves the theorem.
q.e.d.

4 Conclusions

We have provided a model of endogenous party formation in which there are only two parties in equilibrium. These parties propose candidates in extreme positions of the policy space. The adopted policy, however, is a compromise between these two extremes.

The two key assumptions in the model are: (i) the weight function is continuous; (ii) the policy space is one-dimensional and atomless. Our main result does not generalize to the case of a measure space with atoms. The reason is that in the presence of atoms the outcome is not a continuous function of the strategies. Thus, in general, the polarized strategies in the theorem do not form a strong Nash equilibrium. As an example, consider the case of a distribution of voters given by three atoms of equal mass such that

one of them is at point 1/2 and the other two are located at the extremes of the policy space. It is not difficult to see that at the unique Nash equilibrium each type votes for her ideal policy and so there is no polarization of agents strategies. Obviously, if we allow types to use mixed strategies we retain continuity of the outcome function and therefore a smooth partition of the electorate in two parties each offering an extreme proposal. In this case, existence and uniqueness of a polarized SNE follow from exactly the same arguments used in the proof of the theorem.

References

[1] Alesina A, Rosenthal H (1995) Partisan Politics. Cambridge University Press, Cambridge
[2] Le Breton M, Weber S (1995) Stability of Coalition Structures and the Principle of Optimal Partitioning. In: Barnett W, Moulin H, Salles M, Schofield N, (eds) Social Choice, Welfare and Ethics. Cambridge University Press, Cambridge
[3] Fedderson T (1992) A Voting Model Implying Duverger's Law and Positive Turnout. American Journal of Political Science 36: 938 962
[4] Greenberg J, Weber S (1993) Stable Coalition Structures with a Unidimensional Set of Alternatives. Journal of Economic Theory 60: 62 82
[5] Greenberg J, Weber S (1995) Multiparty Equilibrium under Proportional Representation. American Political Science Review 76: 693 703
[6] Halmos P.R. (1951) Measure Theory. D. van Nostrand, New York
[7] Ortuno Ortin I (1997) A Spatial Model of Political Competition and Proportional Representation. Social Choice and Welfare, 14: 427 438