Ideology and Existence of 50%-Majority Equilibria in Multidimensional Spatial Voting Models *

Hervé Crès
Institut d'Études Politiques de Paris†

M. Utku Ünver
Boston College‡

September 17, 2008

Abstract

When aggregating individual preferences through the majority rule in an n-dimensional spatial voting model, the 'worst-case' scenario is a social choice configuration where no political equilibrium exists unless a super majority rate as high as $1 - 1/n$ is adopted. In this paper we assume that a lower d-dimensional ($d < n$) linear map spans the possible candidates' platforms. These d 'ideological' dimensions imply some linkages between the n political issues. We randomize over these linkages and show that there almost surely exists a 50%-majority equilibria in the above worst-case scenario, when n grows to infinity. Moreover the equilibrium is the *mean voter*. The speed of convergence (toward 50%) of the super majority rate guaranteeing existence of equilibrium is computed for $d = 1$ and 2.

Keywords: Spatial voting, super majority, ideology, mean voter theorem, random point set.

JEL-classification: D71 and D72

*All errors are our own responsibility.

†E-mail: herve.cres@sciences-po.fr. Address: Institut d'Études Politiques de Paris, 27, rue Saint-Guillaume 75007 Paris, France.

‡E-mail: unver@bc.edu. Address: Boston College, Department of Economics, 140 Commonwealth Ave., Chestnut Hill, MA, 02467, USA.

1 Introduction

It is well known since Plott (1967) that a 50%-majority stable political equilibrium typically does not exist in a multidimensional voting setup. A way to restore existence of a stable outcome is to require a super majority rule to overrun the status quo, thus giving rise to the concept of ρ-majority equilibrium, where $\rho \in [1/2, 1]$ is the proportion of the voting population a challenger must rally to take over. It is widely admitted that the smaller the rate of super majority needed to secure existence of an equilibrium (i.e., the less conservative the voting rule), the better.

There is a wide literature on the level of super majority required for existence, both in deterministic or probabilistic setup (see, e.g., Ferejohn and Grether (1974), Caplin and Nalebuff (1988, 1991) and Balasko and Crès (1997)). In a standard social choice setup where agents, endowed with continuous and convex preferences, have to choose among political alternatives in a non-empty, compact and convex subset of $\mathbb{R}^n$, Greenberg (1979) shows that a necessary and sufficient condition for the existence of a ρ-majority equilibrium is $\rho \geq \frac{n}{n+1}$.

To show that this bound is tight, Greenberg (1979) constructs a voting configuration where no incumbent is stable with respect to a super majority rule with rate smaller than $\frac{n}{n+1}$. It follows: Take $n+1$ independent points in $\mathbb{R}^n$ and interpret them as the ideal political choices of $n+1$ voters endowed with euclidean preferences. Denote $\mathcal{S}_n$ the n-dimensional simplex generated by the voters' ideal points. Fix an incumbent $x \notin \mathcal{S}_n$; then $s(x) = \text{ argmin } \{\|x - s\|,\ s \in \mathcal{S}_n\}$ is unanimously preferred to x, hence x is not stable under any ρ-majority rule with $\rho < 1$. Now, fix an incumbent $x \in \mathcal{S}_n$; then it is always possible to find a challenger preferred by n out of the $n+1$ voters: indeed, denote $\bar{\mathcal{S}}_n$ the $(n-1)$-dimensional simplex generated by the ideal points of these n voters ($\bar{\mathcal{S}}_n$ is a face of $\mathcal{S}_n$), then one can reconduct the previous argument, and show that $\bar{s}(x) = \text{ argmin } \{\|x - \bar{s}\|,\ \bar{s} \in \bar{\mathcal{S}}_n\}$ is preferred to x by all of these n voters.

This example is thus a 'worst-case' scenario. One easily sees that if the voters' ideal points are taken in a lower dimensional subspace, then the upper bound decreases. But the gain remains small though. And one gets existence of political equilibria for not too conservative voting rules only when the number, n, of political issues is very low. This bound is $\rho = 1/2$ when $n = 1$ (the so-called 'median voter theorem'); $\rho = 2/3$ when $n = 2$; and for $n \geq 3$, then the required rate of super majority must be above 3/4 (and converges the the unanimity criterion when n goes to infinity), a level very rarely observed in practice. Indeed, constitutions or corporate charters build on super majority rates which are very rarely above 70%[1], although the number of political issues at stake in electoral processes

[1] For decisions on issues which are delegated to the European Union, the rate was 72% in the Maastrich Treaty, it was decreased to a mix of 65% –of the States– and 55% –of the population– in the Constitutional Treaty.

is obviously often very large: it is not rare, when reading political platforms proposed by candidates in large elections, to denumerate several dozens of issues[2]. Hence the question: why, if there are so many issues, do we observe so reasonable super majority rates in practice?

A first answer might be that one should not believe in Greenberg's worst-case scenario. A second answer can be found in the Hinich-Ordeshook spatial voting model[3]. According to the latter, there are only a few political dimensions underlying the platforms proposed by the candidates. These few dimensions are claimed to be *ideological*. Ideologies imply linkages[4] between political issues and thus span a lower dimensional linear space (dubbed the 'campaign space' in the sequel) on which the original distribution of voters' ideal points is projected.

The assumption that political platforms are based on ideology stems from the belief that the cleavages between candidates separate along simpler, more predictable lines than the n-dimensional policy space would imply. As Popkin (1994) states it (p. 51): "Ideology is not the mark of sophistication and education, but of uncertainty and lack of ability to connect policies with benefits... Parties use ideologies to highlight critical differences between themselves, and to remind voters of their past successes". This approach has some empirical relevance: Poole and Rosenthal (1991, 1996) show that in the USA, with the exception of the 32nd Congress, two dimensions are always capable of explaining more than 80% and up to 95% of the variation in the votes of elected officials on most issues. The same, Poole and Rosenthal (1997) and McCarty, Poole and Rosenthal (1997) test the Hinich-Ordeshook spatial voting model on post World War II Congressional roll call voting and show that only two dimensions are required to account for most of the votes: the liberal-conservative continuum[5] and the dimension of conflict over race and civil rights.

But one cannot exclude that the number of underlying ideological dimensions be larger than 3. In the political debate in France on the referendum for ratification of the European constitutional treaty during the Spring of 2005, one cannot explain the cleavages between *and within* parties through the traditional left-right dimension. One also needs the now classical ideological dimension 'sovereignist-federalist' to explain the split of the gaullist party; furthermore the possible future entry of Turkey was an element of the debate, and

[2] Everybody in France recalls the '110 propositions' of the candidate François Mitterrand for the presidential election of 1981.

[3] This model was first proposed by Cahoon, Hinich and Ordeshook (1976), Ordeshook (1976) and Hinich and Pollard (1981) and then developed by Enelow and Hinich (1984) and Hinich and Munger (1994).

[4] These linkages formalize the fundamental insight of Converse (1964) according to which ideology (the Conversian 'belief system') interrelates and bundles the political issues: ideology is fundamentally the knowledge of what-goes-whith-what. As Converse (1964) states it: ideology is "...a configuration of ideas and attitudes in which the elements are bound together by some form of constraint" (p. 207).

[5] A judgemental dimension that has been "highly serviceable for simplifying and organizing events in most Western politics for the past century", Converse (1964, p. 214).

religion (the ideological position in the 'laïcity' dimension) was clearly the only way to explain another (orthogonal) split of the gaullist party.

Another type of political debates often builds on more than three underlying dimensions: proxy fights in publicly traded corporations in a context of market failures[6]. The stakes are probably simpler to grasp than in ordinary political debates; moreover, shareholders usually have access to a more measurable and precise information which is easier to aggregate. Yet corporate charters rarely choose rates of super majority beyond 65%.

The answer to the question "why do we observe so reasonable super majority rate in practice?" seems to be: not only the political competition articulates along fewer, simpler and more predictable lines than the n-dimensional policy space would imply, but also one should not believe in Greenberg's worst-case scenario. The present paper goes one step further. Its main contribution is an aggregation theorem that links the two latter arguments: *one should not believe in Greenberg's worst-case scenario* because *the political competition happens in a lower dimensional subspace* spanned by the underlying ideologies. Indeed, if we randomize on the linkages between issues imputed by ideologies, our main result (Theorem 1) states that *the Hinich-Ordeshook approach almost surely transforms Greenberg's worst-case scenario into the best-case scenario of a symmetric distribution of voting characteristics.* And as a consequence we obtain (Theorem 2) a **mean-voter theorem**: the mean voter happens to almost always be the unique 50%-majority equilibrium, when the number of political issues grows large.

The paper is organized as follows: Section 2 introduces the model, first the classical Downsian spatial voting model (Section 2.1), then its Hinich-Ordeshook sophistication (Section 2.2). Then Section 3 states and proves the aggregation theorems. Section 4 computes a lower bound to the speed of convergence of the expected min-max rate toward 50%; the computations give upper bounds on the expected rate of super majority necessary to sustain the mean voter as a political equilibrium, for any number of voters, when the number of underlying ideologies is smaller than 2. Section 5 ends the paper with some concluding comments.

[6] The heterogeneity of the shareholders' opinions can come from imperfect competition, the incompleteness of financial market structure or the presence of externalities. 'Ideological' dimensions in corporate politics can be: the 'philosophy' with respect to debt vs equity, horizontal vs vertical integration, international diversification, expansion vs concentration...

2 The model

2.1 Voters, platforms and the majority rule

The setup to model the electoral process and voting mechanism is the classical Downsian multidimensional spatial voting model (Downs (1957)). There are n measurable criteria of political activity, so that a **political platform** in the policy space can be represented as an n-dimensional vector: $x \in \mathbb{R}^n$. There are m voters in a set $\mathcal{I}$. Each voter is endowed with an **euclidean preference** relation on $\mathbb{R}^n$: agent i, $1 \leq i \leq m$, has a **preferred choice** in the policy space, $x_i \in \mathbb{R}^n$, and his/her utility function over the space of political choices is decreasing with the euclidean distance from his/her preferred choice:

$$\forall x \in \mathbb{R}^n \qquad u_i(x) = -\|x_i - x\|$$

A **society** is a m-tuple $X = (x_i)_{i=1}^m$.

We measure the stability of a political platform in a given society through the Simpson-Kramer approach. Given two political choices $(a, b) \in \mathbb{R}^n \times \mathbb{R}^n$, $\rho(b, a)$ measures the ratio of the electorate that strictly prefers b to a:

$$\rho(b, a) = \frac{\sharp\{i \in \mathcal{I} | u_i(b) > u_i(a)\}}{m}.$$

The **score** of a political choice $a \in \mathbb{R}^n$ is: $\rho(a) = \max_{b \in \mathbb{R}^n} \rho(b, a)$. Clearly, the score of any political choice taken outside the closed convex hull, $\langle X \rangle$, of X will be 1: the challenger b that minimizes the distance between a and $\langle X \rangle$ is unanimously preferred to a. Hence looking for the 'best' status quo, i.e., the ones with lowest score, we can reduce our search to $\langle X \rangle$. The **min-max rate** of society X is: $\rho^* = \min_{a \in \mathbb{R}^n} \rho(a)$. The **min-max set** of society X is: $S^*(X) = \{a \in \mathbb{R}^n | \rho(a) = \rho^*\}$.

The **majority rule** with rate $\rho \in [0, 1]$ states that candidate b is preferred by society X to (or defeats) candidate a if and only if $\rho(b, a) > \rho$. A candidate a is said to be **ρ-majority stable** in society X if and only if there is no alternative that defeats it, i.e., if and only if its score is not larger than ρ: $\rho(a) \leq \rho$. Such a candidate is a **political equilibrium** for the majority rule with rate ρ.

One knows since the seminal work of Plott (1967) that 50%-majority stable equilibria generally do not exist when $n \geq 2$. To recover existence of political equilibria, one has to impose a super majority voting rule, i.e., a voting rule with rate $\rho > 1/2$. This paper deals with existence of such political equilibrium based on super majority voting. Along that search, political platforms in the min-max set have this appealing property that they are equilibria for the lowest rate of super majority, hence the *less conservative* voting rule.

The super majority rate one has to impose in order to recover existence of equilibrium can be quite high, though: As extensively explained in the introduction, suppose that $m = n+1$

and the m-tuple X are the vertices of an n-dimensional simplex, then obviously any political choice in $\langle X \rangle$ has a score of $n/(n+1)$ (it is enough to choose a challenger closer to any of the $n+1$ $(n-1)$-dimensional faces of the simplex). Therefore, since any political choice outside $\langle X \rangle$ has score 1, the min-max rate is $n/(n+1)$ and the min-max set is $\langle X \rangle$: one has to impose a super-majority rule of rate $\rho \geq n/(n+1)$ to get a stable political choice, and then all choices in $\langle X \rangle$ are ρ-majority stable. Greenberg (1979) proves that the condition $\rho \geq n/(n+1)$ —to get existence of a ρ-majority stable political equilibrium— is in fact *sufficient* as soon as the voter's preferences satisfy very mild properties of continuity and convexity. Hence, the latter case is a *worst-case scenario*, as far as getting not too conservative min-max rate is concerned. The present paper can be read as an attempt to downside the relevance of this worst-case scenario.

Some convincing arguments along the same line are available in the social choice literature. One of the most important one is given in Caplin and Nalebuff (1988, 1991). They give a dimension-free upper bound to the min-max rate under the conditions that preferred choices of agents are selected from a σ-concave distribution with compact and convex support. This upper bound (which, asymptotically, is lower than 64%) is given by the score of the *mean voter*, the voter whose preferred choice is the barycenter of all x_i's. This literature can roughly be regarded as looking for multi dimensional versions of the median voter theorem.

2.2 Ideology, candidates and political campaigns

A central assumption of our model is that, although the number (here: n) of criteria for political activity can indeed be quite large, the political competition takes place in a subspace of lower dimension: $d < n$. In accordance with the Hinich-Ordeshook spatial voting model, this lower dimensional space is considered to be the **ideological space**, assumed to be $\mathbb{R}^d$ without loss of generality. According to this approach, the ideologies are linked to the platforms by a linear map, L, from the ideological space to the policy space: a candidate, $\pi^A \in \mathbb{R}^d$, imputes a platform $x^A \in \mathbb{R}^n$ such that $x^A = x^0 + L\pi^A$, where x^0 is the platform of status quo policies. Finally, the d-dimensional affine subspace which is the image of $\mathbb{R}^d$ by L translated by x^0 is called the **campaign space**, $C \subset \mathbb{R}^n$, in the sequel. Before developing the strength of the model, let us illustrate through an example how the linear map L operates.
<u>An illustration</u>: Issues of political activity are often precise and technical; consider two such classical issues like (1) how much of the State's budget, x_1, must be allocated to buy helicopters, and (2) how much of the State's budget, x_2, must be allocated to create more slots in kindergartens. For the sake of simplicity, we limit the issues to these two, hence $n = 2$. The assumption is made that platforms proposed by candidates in this two-dimensional policy space can be explained through a (say) one-dimensional underlying linear subspace,

e.g., the classical liberal-conservative (left-right) dimension; hence $d = 1$. Given a vector of status quo policies x^0, the sensitivity of x_j, $j = 1, 2$, to the position π of the candidate in the ideological space is a fixed scalar $l_j \in \mathbb{R}$, therefore x_j is an *affine* function of π:

$$x_j = x_j^0 + l_j \pi \ , \quad j = 1, 2 \ .$$

Figure 1 (resp. 2) plots the relation between ideology and helicopters (resp. kindergartens). E.g., the policy regarding kindergarten is almost not sensitive to ideology, and only slightly decreases (l_2 is small and negative) with π: a leftist candidate wants to create slots in kindergartens because these structures are more used by low-class workers than by wealthy families; a rightist candidate uses slots' creation in kindergartens as an incentive to increase fertility. The policy regarding helicopter is more sensitive to ideology, and increasing (l_1 is positive): rightist candidate are usually more hawkish, and spending on helicopters rises as ideology moves right, as shown by the plain line L_1.

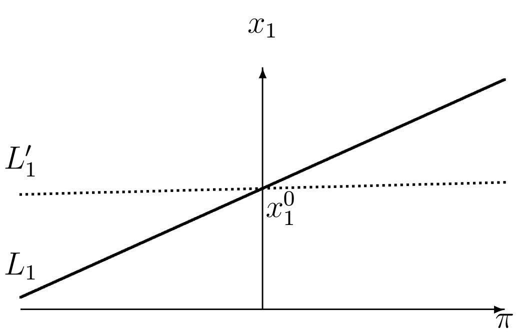

Figure 1

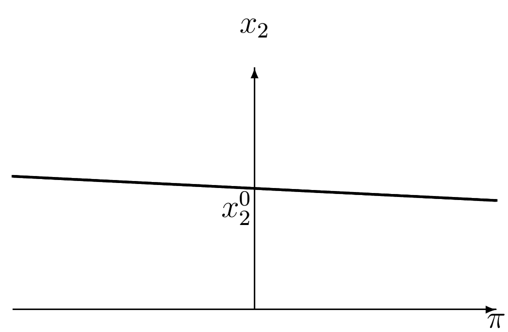

Figure 2

The sensitivity of policies to ideology as depicted on Figures 1 and 2 implies a linkage between the two issues in the policy space: the induced campaign space, C (plain line on Figure 3), is going through the status quo x^0 with slope l_2/l_1. The induced 'ideal candidate' $\bar{x}_i$ of voter i, whose preferred platform is x_i, obtains by orthogonal projection of x_i on the campaign space. And consequently, voter i votes for the candidate whose imputed platform is closest to his 'ideal candidate' $\bar{x}_i$. In the general (n, d) case, the euclidean structure of the original voting configuration gives rise, through the orthogonal projection on C, to a social choice problem involving m voters with euclidean preferences in $\mathbb{R}^d$. Hence we are dealing with a d-dimensional spatial voting problem with m voters and thus we are left with a combinatorial problem about m-tuples of points in $\mathbb{R}^d$ rather than in $\mathbb{R}^n$.

The assumption that political platforms are based on ideology stems from the belief that the cleavages between candidates separate along simpler, more predictable lines than the n-dimensional policy space would imply. Simplicity and predictability makes the voter's

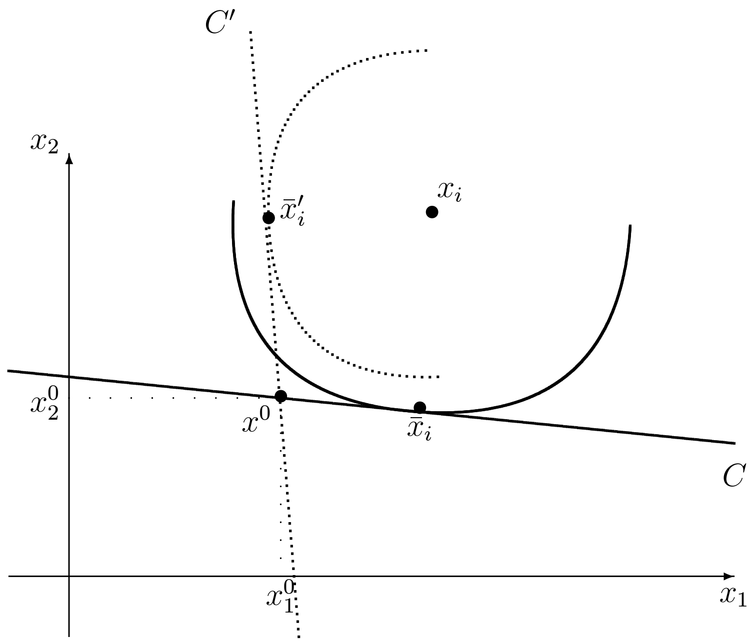

Figure 3

duty of voting easier. Not only ideology transmits information to voters, but also it creates enthusiasm for political action. This is a virtue of this approach on the candidate's side. Of course, reducing the dimensionality of the campaign space also makes communication easier and less costly for candidates. But on top of that, the credibility of its commitment is stronger when his/her actions are perceived by the voters to be based on ideology[7] (see Enelow and Hinich (1984) or Hinich and Munger (1997) for development of these arguments). Last but not least, as underlined in the introduction, the empirical relevance of this approach has been underlined by Poole and Rosenthal (1991, 1996, 1997) and McCarty, Poole and Rosenthal (1997).

If the Hinich-Ordeshook spatial voting model has the virtue of offering a more realistic view of electoral competition, we argue in the present paper that it moreover has *extremely nice properties as far as aggregation of individual preferences* is concerned. Indeed, we prove in the sequel that the worst-case configuration of the society (worst case as far as aggregation is concerned) –i.e., when the point-set X is an $(n+1)$-tuple of points forming a n-dimensional simplex with equal voting rights on the vertices– transforms *almost surely* into a best-case configuration –i.e., the $(n+1)$-tuples of projected points in $\mathbb{R}^d$ X is symmetrically distributed.

The first step of our argument is to qualify what we mean by 'almost surely'. Let us go

[7] As Hinich and Munger puts it, "classical spatial theory assumes that each politician chooses the position that maximizes his or her vote share given the expected platform of the opponents. (...) Such an approach (...) may be of little use in describing real world politics. For such an approach to work, voters must believe that a candidate who takes a position is likely to deliver that position. (...) The candidate's promise must be 'credible'.

back to the above illustration. Suppose now that an exogenous historical shock occurs, e.g., a terrorist attack. Most probably this event is going to impact the sensitivity of the first issue (helicopters) to ideology: all candidates become hawkish and want to invest more into such a modern defense tool as helicopters, independently of his/her ideology. Hence a new line L'_1, with a much smaller sensitivity rate: $l'_1 < l_1$ (see the —almost flat— dotted line on Figure 1). It is probably going to be the case that everybody in the society, candidates *and voters* are going to prefer an absolute increase $\Delta x_1 > 0$ in the political platform; we assume that in such an event the perturbed status quo becomes: $x_1^{'0} = x_1^0 + \Delta x_1$, and that for all i, the preferred platform's first component becomes: $x'_{i1} = x_{i1} + \Delta x_1$. Hence this general absolute increase Δx_1 results in a *global (rightward) translation* of the spatial point-set configuration, and this translation has no impact on the geometric properties of our problem. Therefore, without loss of generality, we can consider Δx_1 to be zero, and the only impact of this exogenous historical event is a drop in the sensitivity rate l_1. This results into a new campaign space C' (dotted on Figure 3) going through the status quo x^0 with slope l_2/l'_1. The new induced 'ideal candidate' $\bar{x}'_i$ of voter i obtains by orthogonal projection of x_i on the new campaign space C'.

The idea is that such *random shocks* always happen, although fortunately not all as dramatic as a terrorist attack, and that their 'media' treatment and destiny can change the sensitivities of various issues to ideology. Then the central question is: How is this d-dimensional campaign space chosen? In the present paper, we take a purely Laplacian perspective and assume that C is selected *at random*, according to a 'uniform' distribution on the natural underlying space. We define C as an element in the Grassmanian $G(n, d)$ of oriented d-subspaces in $\mathbb{R}^n$. Random historical and mediatic shocks generate a probability distribution over $G(n, d)$. Among the latter ones, one arises 'naturally': the unique rotation-invariant probability measure, $\mu(n, d)$ (known as the Haar probability measure), on $G(n, d)$, which intuitively selects all d-dimensional campaign spaces 'with equal probability'. Hence $\mu(n, d)$ will be dubbed **impartial** in the sequel. The idea behind impartiality is that the main themes at stake in a political campaign depend heavily on the exogenous shocks of recent history, and the exogenous treatment by the media of these shocks.

3 Main result

For any selected campaign space C, the original social choice problem characterized by the point set X in $\mathbb{R}^n$ gives rise to a lower dimensional social choice problem characterized by the (orthogonally projected) point set $\overline{X}$ in $\mathbb{R}^d$. Suppose X is a $m-1$ dimensional simplex in $\mathbb{R}^n$, such that for each i, x_i (a column vector in $\mathbb{R}^n$) is the i^{th} vertex of simplex X. We say that X is regular if $\|x_i - x_j\| = \|x_i - x_k\|$ for any i, j, k. The simplex is O-centered if

$\sum_i x_i = 0$. Note that for an O-centered regular simplex, we have $x_i^T x_j = x_i^T x_k < 0$ for any i, j, k. In this section, we mute the translational part of the political shock as explained in the previous section and focus on the rotational shock to the campaign space. First, we will consider rotations pivoted at the center of gravity of the simplex X, that is, when the mean-voter is the status-quo (i.e. it is always on the campaign space). Center of gravity is normalized to O, center of the coordinate system. As we will show, this has no cost in our approach, since the center of gravity is the unique 50%-political equilibrium (Theorem 2).

Theorem 1 *Let the campaign space C be impartially randomly selected. The point set $\overline{X}$ coincides in distribution with a negatively correlated sample from a symmetric probability distribution in $\mathbb{R}^d$ which becomes asymptotically independent as $n \to \infty$ with the rate $\frac{1}{n}$ when X is a regular O-centered simplex.*

Proof of Theorem 1: Let $n \geq m > d$. Take a regular $(m-1)$-dimensional O-centered simplex in $\mathbb{R}^n$ whose vertices are column vectors of an $n \times m$ matrix $X = [x_1 \ x_2 \ ... \ x_m]$.[8] We have $x_i^T x_j < 0$ for any i and j. Let Π be $d \times n$ with $\pi_{ii} = 1$ for $i \in \{1, 2, ..., n\}$ and all other entries of Π are 0. We can denote the random $d-$subspace by $\mathbf{C} = \Pi \mathbf{R}$ where $\mathbf{R}$ is a random rotation matrix distributed with the Haar probability measure among the $n \times n$ rotation matrix group denoted by $\mathcal{R}(n)$, that is, every $n \times n$ rotation matrix is chosen with equal probability as a draw of $\mathbf{R}$. Note that $\mathbf{C}$ is distributed with probability measure $\mu(n, d)$ in Grassmanian $G(n, d)$. Note that every rotation matrix is an orthogonal matrix. Let $\mathcal{O}(n)$ denote the $n \times n$ orthogonal matrix group. First, we will consider orthogonal matrices instead of rotation matrices. Let $\mathbf{C}^* = \Pi \mathbf{A}$ be such that $\mathbf{A}$ is a random orthogonal matrix distributed with the Haar probability measure in $\mathcal{O}(n)$, that is, every $n \times n$ orthogonal matrix is chosen with the same probability as a draw of $\mathbf{A}$. Note that orthogonal transformations include rotoinversions (where $\det(A) = -1$) and rotations ($\det(A) = 1$), we will rule out rotoinversions later. First note that every column of $\mathbf{A}$ has a symmetric distribution and $E\mathbf{A} = 0$, since if A is orthogonal then $-A$ is orthogonal and both $\mathbf{A} = A$ and $\mathbf{A} = -A$ are equally likely events. Since orthogonal transformation preserves the inner-product of two vectors, we have $\sum_i a_{ij} a_{ik} = 0$ for $j \neq k$ and $\sum_i a_{ij}^2 = 1$ for any orthogonal matrix A. Since every row permutation and column permutation of A is equally likely to occur, $E\left(\mathbf{a}_{ij}^2\right)$ and $E(\mathbf{a}_{ij}\mathbf{a}_{k\ell})$ is constant for every i, j, k and ℓ. First, $\sum_i E\left(\mathbf{a}_{ij}^2\right) = 1$ implies $E\left(\mathbf{a}_{ij}^2\right) = \frac{1}{n}$. Moreover, $\sum_i E(\mathbf{a}_{ij}\mathbf{a}_{ik}) = 0$ implies $E(\mathbf{a}_{ij}\mathbf{a}_{ik}) = 0$, implying with the symmetry argument that $E(\mathbf{a}_{ij}\mathbf{a}_{k\ell}) = 0$ for every i, j, k, ℓ such that $i \neq k$ or $j \neq \ell$. We are interested in the distribution of the columns of $\mathbf{P} = \mathbf{C}^* X = \Pi \mathbf{A} X$, the projection of X to the random subspace $\mathbf{C}^*$. We will show that each column vector has identical symmetric distribution and each pair of column vectors are negatively correlated. The two events $\mathbf{P} = \Pi A X$ and $\mathbf{P} = \Pi(-A) X$

[8] Let $(i, j)^{\text{th}}$ entry of a matrix H be denoted by h_{ij} and j^{th} column vector of H be denoted by h_j.

are equally like to occur, therefore each column of $\mathbf{P}$ has a symmetric distribution. Since each x_i has the same length, each column vector $\mathbf{A}x_i$ is an identical random vector. Let $\Sigma^{ij} = cov(\mathbf{A}x_i, \mathbf{A}x_j)$ for $i \neq j$. The $(g,h)^{\text{th}}$ entry of Σ^{ij} is $\sigma^{ij}_{gh} = E\left(\left(\sum_k \mathbf{a}_{gk}x_{ki}\right)\left(\sum_\ell \mathbf{a}_{h\ell}x_{\ell j}\right)\right)$. For $g \neq h$,

$$\sigma^{ij}_{gh} = \sum_k \sum_\ell \underbrace{E\left(\mathbf{a}_{gk}\mathbf{a}_{h\ell}\right)}_{=0} x_{ki}x_{\ell j} = 0$$

For $g = h$,

$$\sigma^{ij}_{hh} = \sum_k \sum_\ell E\left(\mathbf{a}_{hk}\mathbf{a}_{h\ell}\right) x_{ki}x_{\ell j} = \sum_k \underbrace{E\left(\mathbf{a}^2_{hk}\right)}_{=\frac{1}{n}} x_{ki}x_{kj} + \sum_k \sum_{\ell \neq k} \underbrace{E\left(\mathbf{a}_{hk}\mathbf{a}_{h\ell}\right)}_{=0} x_{ki}x_{\ell j} = \frac{1}{n}x_i^T x_j < 0$$

Since $\mathbf{p}_i$ is the first d coordinates of $\mathbf{A}x_i$ for each i, different coordinates of each pair of $\mathbf{p}_i$ and $\mathbf{p}_j$ are independently sampled and the same coordinates of each pair $\mathbf{p}_i$ and $\mathbf{p}_j$ are negatively correlated, where correlation goes to zero as $n \to \infty$ (or $m \to \infty$ since $m \leq n$). We conclude our proof by observing that by Theorem 1 of Baryshnikov and Vitale (1994), above matrix $\mathbf{A}$ can be swapped with a random rotation matrix $\mathbf{R}$, and point set $\overline{X}$ consists of draws of columns of $\mathbf{P} = \mathbf{C}X$. ■

Our next result states that there is a unique political equilibrium at the mean voter for 50%-majority rule as the number of voters goes to infinity.

Theorem 2 ***(Mean Voter Theorem)*** *Fix d. Take $m \to \infty$ then almost surely O, the mean voter, is the unique political equilibrium for the 50%-majority rule.*

Proof of Theorem 2: Let $f : \mathbb{R}^d \to \mathbb{R}_+$ be the underlying limiting marginal probability density function (p.d.f.) for the columns of $\lim_{m\to\infty} \overline{X}$. Let $E\rho(w)$ denote the expected value of the score of any point $w \in \mathbb{R}^n$.

First, we show that $O\,(\in \mathbb{R}^n)$, the mean voter, is a political equilibrium for the 50% majority rule. Negatively correlated sampling of political positions has a smaller min-max rate than independent sampling of political positions from the same symmetric marginal distribution. Let $\overline{Y}$ be a set of points independently sampled from p.d.f. f. Under independent sampling, by Theorem 3 in Caplin and Nalebuff (1988) the min-max rate of O for $\overline{Y}$ converges almost surely to the min-max rate of f at $O \in \mathbb{R}^d$ when $m \to \infty$. The min-max rate of O for f is 0.5, since f is symmetric around $O\left(\in \mathbb{R}^d\right)$ by Theorem 1. Since min-max rate cannot smaller than 0.5, min-max rate of O for $\overline{X}$ also almost surely converges to 0.5, that is, $\lim_{m\to\infty} E\rho(O) = 0.5$, implying $O\,(\in \mathbb{R}^n)$, the mean voter, is an equilibrium point for the 50%-majority rule.

Next, we prove that $O\,(\in \mathbb{R}^n)$ is the unique equilibrium issue. Take any issue vector $w \in \mathbb{R}^n \backslash \{O\}$ when m and $n \to \infty$. For finite n, let the first n entries of w be relevant. We will show that $\lim_{m\to\infty} E\rho(w) > 0.5$. Let $\mathbf{C} = \Pi\mathbf{R}$ be the random campaign space spanned by the Haar measure, like in the proof of Theorem 1 where $\mathbf{R}$ is the $n \times n$ impartial random rotation matrix distributed with Haar measure on $\mathcal{R}(n)$ and Π is the $d \times n$ matrix with $\pi_{ii} = 1$ for $i \leq d$ and all other entries of Π are zero. The random projection of the issue vector w on the campaign space is $\Pi\mathbf{R}w$. Note that $\Pi\mathbf{R}w = O \in \mathbb{R}^d$ with probability 0 and $\Pi\mathbf{R}w \neq O \in \mathbb{R}^d$ with probability 1, since $w \neq O \in \mathbb{R}^n$. Hence, the only relevant draws of $\mathbf{R}$ for the calculation of $E\rho(w)$ are all $R \in \mathcal{R}(n)$ such that $\Pi Rw \neq O \in \mathbb{R}^d$. Fix a rotation matrix $R \in \mathcal{R}(n)$ such that $\Pi Rw \neq O$. Since f is symmetric around $O \in \mathbb{R}^d$ by Theorem 1, the min-max rate of the point ΠRw for f is greater than 0.5. Hence expected min-max rate of the projection $\Pi\mathbf{R}w$ is greater than 0.5 when $m \to \infty$, implying $\lim_{m\to\infty} E\rho(w) > 0.5$ and concluding that w cannot be stable under 50% majority voting rule. ■

This result is inspired from a modern approach (in the literature from discrete and computational geometry) to generate random points. When dealing with a d-dimensional spatial voting problem with m voters, one is left with a combinatorial problem about m-tuples of (random) points in $\mathbb{R}^d$. Many 'natural' distributions of these random points have been proposed in the mathematical literature (see Schneider (2004)). Among them, the one described above takes a central place: Every configuration of $m > d$ numbered points in general position in $\mathbb{R}^d$ is affinely equivalent to the orthogonal projection of the set of numbered vertices of a fixed regular $(m-1)$-dimensional simplex onto a unique d-dimensional linear subspace in $\mathbb{R}^n$. This construction builds a one-to-one correspondence between the (orientation-preserving) affine equivalence classes of such point set configurations and an open dense subset of the Grassmanian $G(n, d)$ of oriented d-spaces in $\mathbb{R}^n$. The so-called *Grassmann approach* (sometimes referred as the Goodman-Pollack model) considers the probability distribution on the set of affine equivalence classes of m-tuples in general position in $\mathbb{R}^d$ that stems from the unique rotation-invariant probability measure on $G(n, d)$. Baryshnikov and Vitale (1994) (following an observation of Affentranger and Schneider (1992)) proved that under the Grassmann approach, the resulting point set coincides in distribution with a *standard Gaussian sample* in that subspace. As a consequence, an affine-invariant functional of m-tuples with this distribution is stochastically equivalent to the same functional taken at an i.i.d. m-tuple of standard normal points in $\mathbb{R}^d$.

Our model can be seen as giving another interpretation to the Grassmann approach: We do not use it as a random generation of social choice configurations (m-tuples of points in $\mathbb{R}^d$) according to a 'natural' probability distribution, but as a random generation of a lower dimensional campaign spaces for any original (higher dimensional) social choice configuration. The latter in particular can be a worst-case scenario as defined in the introduction and Section 2.

Remark: A direct consequence of the Grassmann approach is that if we depart from a (regular) worst-case scenario in $\mathbb{R}^n$ and take X to be the vertices of a *regular* $(n-1)$-dimensional simplex (which can as well be translated to be spherico-regular), then, up to an affine transformation, the resulting point set $\overline{X}$ coincides in distribution with a standard Gaussian sample in $\mathbb{R}^d$. In our approach, we do not need linear images and affine translations, since economically they do not correspond to any interpretations, and mathematically we do not require stochastic independence and an underlying Gaussian distribution.

4 Low dimensional social choice configurations

Of course, the asymptotic result in Theorem 2 does not give much idea of the rate of super majority, $\rho(O)$, that one should impose in order to have existence of a political equilibrium when d and m are 'small', nor does it give an idea of its *expectation*, $E\rho(O)$. In this section, we give the an *upper-bound* to *distribution-free* value of this expected score for $d=1$ and $d=2$. We consider *independent sampling* of m points drawn from a symmetric distribution about O ($\in \mathbb{R}^d$ for $d=1$ and $d=2$) which approximate the sampling for $n \to \infty$. In our approach in Theorem 1, we need *negatively-correlated sampling* for finite n. Therefore, the results derived in this section are upper bounds to actual $E\rho(O)$ for finite n when $d=1$ and $d=2$ for any n with X is regular and O-centered. Let $n \to \infty$ in this section. For uni- and bidimensional campaigns, these upper-bounds that we derive are plotted at the end of the section in Figure 7.

Unidimensional campaigns

Proposition 1 *Set $d=1$. Consider an m-sample independently drawn from a distribution which is symmetric about O, then the expected value of the score of O is*

$$E\rho(O) = \frac{1}{2} + \frac{1}{2^{2\left[\frac{m}{2}\right]+1}} \binom{2\left[\frac{m}{2}\right]}{\left[\frac{m}{2}\right]}$$

where $[x]$ denotes the integer part of x.

Proof of Proposition 1: Consider the following process which is due to Wendel (1962): choose m random points in an interval centered at O: $q_1, q_2, \ldots, q_m$. For each i, $1 \leq i \leq m$, set p_i equal to q_i or to $-q_i$ with equal probability 1/2. The points $p_1, \ldots, p_m$ are again i.i.d. random points in the interval (we ignore the degenerate configurations, occurring with probability zero, where two points q_i and q_j would be equal or opposed).

Consider first the case of an even m: $m=2p$. The score of O will be the highest ratio of points on either side of O. Obviously if there are k points on one side (and $m-k$ on the other side), $0 \le k \le p$, this number is $(m-k)/m$, and this happens with probability $2\frac{1}{2^m}\binom{m}{k}$ when $0 \le k \le p-1$, and with probability $\frac{1}{2^m}\binom{m}{k}$ when $k = p$. Hence

$$\begin{aligned} E\rho(O) &= \frac{1}{2^m}\binom{2p}{p}\frac{p}{2p} + \frac{1}{2^{m-1}}\sum_{k=0}^{p-1}\binom{m}{k}\frac{m-k}{m} \\ &= \frac{1}{2^{m+1}}\binom{2p}{p} + \frac{1}{2^{m-1}}\sum_{k=0}^{p-1}\binom{m-1}{k} \end{aligned}$$

and the second term on the right-hand side is $1/2$. Now if $m = 2p+1$:

$$\begin{aligned} E\rho(O) &= \frac{1}{2^{m-1}}\sum_{k=0}^{p}\binom{m}{k}\frac{m-k}{m} = \frac{1}{2^{m-1}}\sum_{k=0}^{p}\binom{m-1}{k} \\ &= \underbrace{\frac{1}{2^{m-1}}\sum_{k=0}^{p-1}\binom{m-1}{k} + \frac{1}{2^m}\binom{m-1}{p}}_{=\frac{1}{2}} + \frac{1}{2^m}\binom{m-1}{p} \end{aligned}$$

All we used in this proof was that the original distribution is symmetric about O and that some degeneracies occur with probability zero. Hence the result, which is distribution-free. ■

Bidimensional campaigns

We consider the same process as in Proposition 1 and choose m random points in a disk centered at O: $Q_1, Q_2, \ldots, Q_m$. For each i, $1 \le i \le m$, we set P_i equal to Q_i or to $-Q_i$ with equal probability $1/2$ (without loss of generality, we can choose the Q_i's on the same side of a hyperplane through 0 as in Figure 4 below; Figure 5 corresponds to the configuration: $P_i = Q_i$ for $i = 1, 3, 4$ and $P_i = -Q_i$ for $i = 2, 5$). The points $P_1, \ldots, P_m$ are again i.i.d. random points in the disk. The original question answered by Wendel (1962), see also Wagner and Welzl (2001), was: what is the probability that O is not in the convex hull of the P_i's? (In other words, what is the probability that the score of O be 1?) The answer is: $m/2^{m-1}$. Indeed, independently of the choice of the Q_i's (again, we ignore the degenerate configurations, occurring with probability zero, where two vectors Q_i and Q_j would be collinear), there are $2m$ possibilities to choose the signs of the P_i's such that O can be separated from these points by a line (every partition of the Q_i's by a line through O gives two such possibilities). Again, all we used in this line of reasoning was that the original distribution is symmetric about O and that some degeneracies occur with probability zero. Hence the result is, once more, distribution-free.

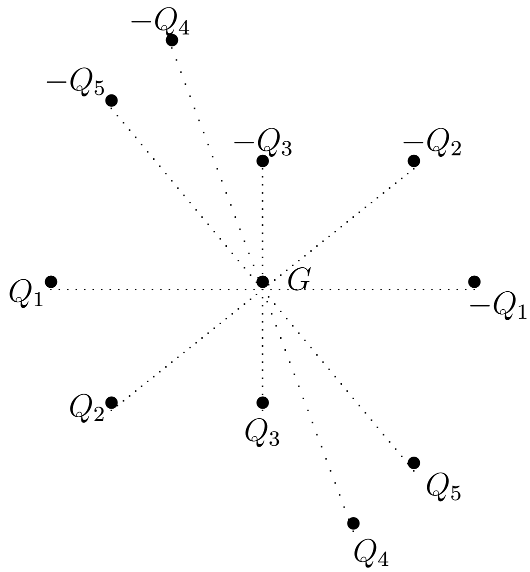

Figure 4

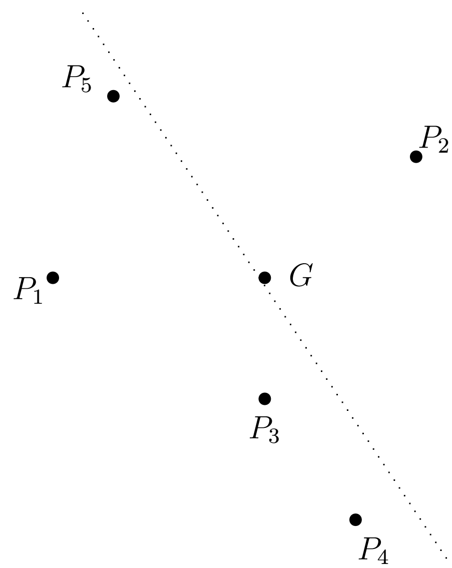

Figure 5

We now proceed along this line of reasoning and compute the probability that the score of O be j/m. E.g., in the social choice configuration shown on Figure 5, the score of O is 4/5 as shown by the dotted separation line.

Each social choice configuration can be described as a (q, p)-**sequence** (or random walk) of plus ones and minus ones, according to whether P_i is equal to Q_i (then +) or not (then -), $1 \leq i \leq m$: $\epsilon_1, \ldots, \epsilon_m$, with, say, q plus ones and p minus ones, $q+p=m$. The configuration on Figure 5 corresponds to the (3,2)-sequence: $+-++-$. The partial sum $s_k = \epsilon_1 + \ldots + \epsilon_k$ represents the difference between the number of pluses and minuses occurring at the first k places, $0 \leq k \leq m$, with $s_0 = 0$ and $s_m = q - p$. Define: $\bar{s} = \max_k s_k$ and $\underline{s} = \min_k s_k$.

Lemma 1 *In the social choice configuration represented by a (q, p)-sequence $(\epsilon_1, \ldots, \epsilon_m)$, the score of O is* $\rho(O) = \dfrac{\max\{q - \underline{s}, p + \bar{s}\}}{m}$.

Proof of Lemma 1: Consider a line which separates the $+Q_i$'s from the $-Q_i$'s and passing through O. It has q of the P_i's on one side and p on the other side. Now turn this line by pivoting at O so that it goes in-between Q_1 and Q_2: it has now $q - s_1$ of the P_i's on one side and $p + s_1$ on the other side. Now turn it by pivoting at O so that it goes in-between Q_2 and Q_3: it has now $q - s_2$ of the P_i's on one side and $p + s_2$ on the other side. And so on. The maximum number of P_i's on one side of a line through O is therefore $\max\{\ldots, q - s_k, \ldots, p + s_k, \ldots\} = \max\{q - \underline{s}, p + \bar{s}\}$. Hence the result. □

To compute the probability that the score of O be j/m, we need to compute the number of (q,p)-sequences such that $\max\{q-\underline{s}, p+\bar{s}\} = j$. To do that, we follow a classical geometric method in the standard orthonormal basis where the x-axis in horizontal and the y-axis is vertical. Following Feller (1968), the sequence $(\epsilon_1, \ldots, \epsilon_m)$ is identified with a **path** from the origin to the point $(m, q-p)$: this path is a polygonal line whose vertices have abscissa $0, 1, \ldots, m$ and ordinates $s_0, s_1, \ldots, s_m = q-p$; $\bar{s}$ is the highest point of the path and $\underline{s}$ the lowest. Obviously there are[9] $\binom{m}{p}$ such paths from the origin to the point $(m, q-p)$: as many as there are ways of choosing the p places for the minuses out of the m possibilities.

Note that for a (q,p)-sequence, $\bar{s} \geq \max\{0, q-p\}$ and $\underline{s} \leq \min\{0, q-p\}$ entail that $\max\{q-\underline{s}, p+\bar{s}\} \geq \max\{q,p\}$, therefore we restrict attention to $j \geq \max\{q,p\}$. And in the case when m is even and $p = q = m/2$, obviously $\max\{q-\underline{s}, p+\bar{s}\} \geq m/2+1$ therefore we restrict attention in that case to $j \geq m/2+1$. Hence we consider j such that $[m/2]+1 \leq j \leq m$.

A (q,p)-sequence is such that $\max\{q-\underline{s}, p+\bar{s}\} = j$ whenever the associated path remains in the corridor between the lines $y = j-p$ and $y = q-j$, and hits at least one of them (see Figure 6 drawn for the configuration of Figure 5 and $j = 4$).

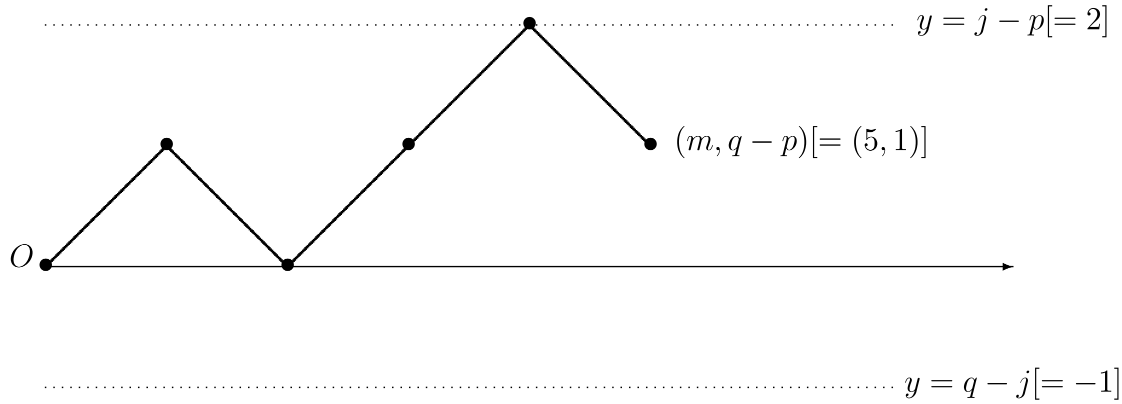

Figure 6

Lemma 2 *Fix j, $[m/2]+1 \leq j \leq m$. The number of (q,p)-paths such that* $\max\{q-\underline{s}, p+\bar{s}\} = j$ *is*

$$a_{m,q,j} = \begin{cases} A_{m,q,j+1} - A_{m,q,j} & \text{if } m-j \leq q \leq j \\ 0 & \text{otherwise} \end{cases}$$

where for $m-j \leq q \leq j$

$$A_{m,q,j} = \sum_k \left[\binom{m}{q+k(2j-m)} - \binom{m}{j+k(2j-m)} \right] \tag{1}$$

[9] By convention, the combination number will be set to zero in case $p < 0$ or $p > m$.

(the series extending over all integers k from $-\infty$ to $+\infty$, but having only finitely many non-zero terms) is the number of (q,p)-paths such that $\max\{q-\underline{s}, p+\bar{s}\} < j$*: those which hit neither $y = j - p$ nor $y = q - j$, and the number of (q,p) paths such that* $\max\{q-\underline{s}, p+\bar{s}\} < m+1$ *is given by*

$$A_{m,q,m+1} = \binom{m}{q}. \tag{2}$$

Proof of Lemma 2: The equation relating the $a_{m,q,j}$'s to the $A_{m,q,j}$'s is immediate. The computation of the $A_{m,q,j}$'s relies on repeated applications of the 'reflection principle' due to Désiré André (see, e.g., Feller (1968), Chapter III) and is done in the appendix. ■

Proposition 2 *Consider an m-sample independently drawn from distribution which is symmetric about O, then the probability that the score of O be expected value of the score of O be j/m, $[m/2]+1 \leq j \leq m$ is*

$$\bar{a}_{m,j} = \frac{1}{2^m} \sum_{q=m-j}^{j} a_{m,q,j}$$

and the expected value of the score of O is

$$E\rho(O) = \sum_{j=[m/2]+1}^{m} \frac{j\bar{a}_{m,j}}{m} .$$

Proof of Proposition 2: Immediately follows from Lemma 2. ■

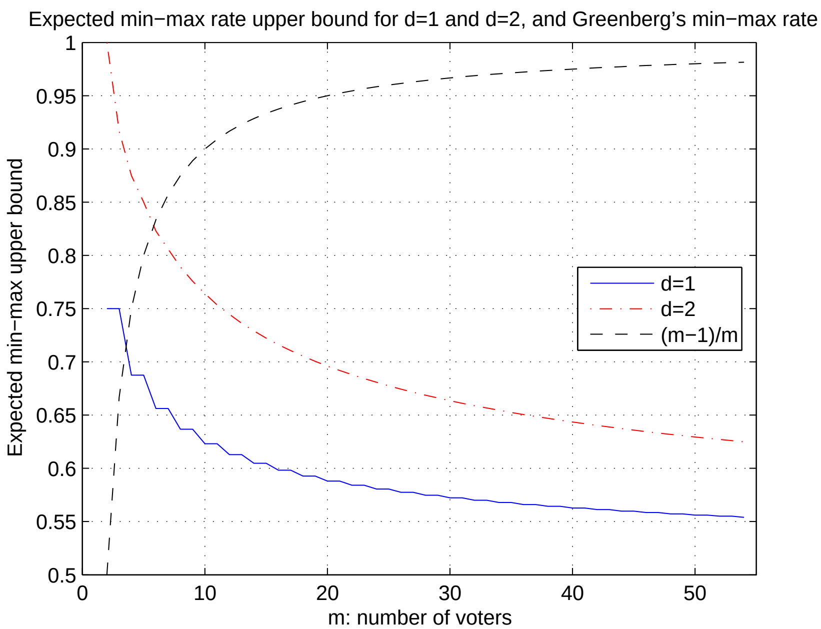

Figure 7: The min-max rate for $n \to \infty$ when $d = 1$ and $d = 2$. For finite n, these figures are upper bounds for the actual min-max rates.

5 Concluding Comments

The present paper proposes a theorem of aggregation of individual preferences through the 50% majority rule in a multidimensional spatial voting model. Of course, the result is obtained at a non-negligible cost in terms of assumptions: first the regularity of the simplicial distribution of voters' ideal points; second the 'uniform distribution' on the set of linkages between issues imputed by ideologies (the Haar probability measure on $G(n, d)$). Of course, the robustness of the results when one relaxes these two assumptions should be studied. But on the other hand it is quite strong since it gives existence for the 50% majority rule. Another important aspect is that it fingers the mean voter as the candidate most likely to be stable in the voting process. Mean voter theorems are very welcome in public economics because in many contexts the mean voter is the one who has the right incentives as far as making an economically efficient choice is concerned.

We would like to stress one last point. An important property of the approach chosen here is that it is compatible with the idea that politicians "die in their ideological boots".

Poole (2003) shows a variety of evidence that members of the US Congress are ideologically consistent: they adopt an ideological position and maintain it over time. An interpretation of Theorem 2 is that *in the long run*, ignoring the historical, sociological and mediatic shocks which are going to shape the linkages between political issues, this might be a good strategy not to change one's mind! A strategic politician should choose an ideological position that he/she believes will place, as frequently as possible over the years, his/her imputed platform at the center of gravity of the voters' ideal points. Different ideological positions come from different tastes, but also from different priors on the distribution of historical shocks (the so-called 'sens de l'histoire') and therefore of linkages between issues. Maintaining that ideological position over time is essential for their credibility, and thus an important asset for future political successes. Now, what is a good strategy *in the short-run*? A strategy here is neither the choice of an ideological position (basically chosen once for all at the beginning of one's career —although there might be more than one beginning...) nor the choice of a political platform (automatically imputed by the ideological position), but an action that impacts the linkages between issues in a way that places the candidate at the center of gravity of the projected set of voters' ideal points. This is left for future work.

References

Affentranger, F. and Schneider, R. (1992). Random projection of regular simplices, *Discrete Computational Geometry*, Vol. 11, 141-147.

Balasko, Y. and Crès, H. (1997). Condorcet cycles and super-majority rules, *Journal of Economic Theory*, Vol. 70, 437-470.

Baryshnikov, Y.M. and Vitale, R.A. (1994). Regular simplices and Gaussian samples, *Discrete Computational Geometry*, Vol. 7, 219-226.

Cahoon, L., Hinich, M. and Ordeshook, P. (1976). A multidimensional statistical procedure for spatial analysis, manuscript, Carnegie Mellon University.

Caplin, A. and Nalebuff, B. (1988). On the 64%-majority rule, *Econometrica*, Vol. 56, No. 4, 787-814.

Caplin, A. and Nalebuff, B. (1991). Aggregation and social choice: a mean voter theorem, *Econometrica*, Vol. 59, No. 1, 1-23.

Converse, P. (1964). The nature of belief systems in mass publics, in *Ideology and discontent*, D. Apter, ed., New York: Free Press.

Enelow, J. and Hinich, M. (1984). *The spatial theory of voting*, New York: Cambridge University Press.

Feller, W. (1968). *An introduction to probability theory and its applications*, John Wiley and Sons, New York.

Ferejohn, J.A., and Grether, D.M. (1974). On a class of rational decision procedures, *Journal of Economic Theory*, Vol. 8, 471-482.

Grandmont, J.-M. (1978). Intermediate preferences and the majority rule, *Econometrica*, Vol. 46, 317-330.

Greenberg, J. (1979). Consistent majority rules over compact sets of alternatives, *Econometrica*, Vol. 47, 627-636.

Hinich, M. and Munger, M. (1997). *Analytical politics*, New York: Cambridge University Press.

Hinich, M. and Pollard, W. (1981). A new approach to the spatial theory of electoral competition, *American Journal of Political Science*, Vol. 25, 323-341.

McCarty, N., Poole K. and Rosenthal H. (1997). Income Redistribution and the Realignment of American Politics. Washington D.C.: American Enterprise Institute.

Ordeshook, P. (1976). The spatial theory of elections: a review and a critique, in *Party identification and beyond*, I Budge, I. Crewe and D. Farlie, eds., New York: Wiley.

Plott, C. (1967). A notion of equilibrium and its possibility under the majority rule, *American Economic Review*, Vol. 57, 787-806.

Poole, K. (2003). Changing minds? Not in Congress!, manuscript, University of Houston.

Poole, K. and Rosenthal, H. (1991). Patterns of congressional voting, *American Journal of Political Science*, Vol. 35, 228-278.

Poole, K. and Rosenthal, H. (1996). *Congress: a political-economic history of roll-call voting*, New York: Oxford University Press.

Popkin, S. (1994). *The reasoning voter: Communication and persuasion in presidential campaigns* (2nd. edition), Chicago: University of Chicago Press.

Schneider, R. (2004). Discrete aspects of stochastic geometry, *Handbook of Discrete and Computational Geometry*, J.E. Goodman and J. O'Rourke, eds., 2nd ed., Chapman and Hall/CRC, Boca Raton, 255-278.

Wagner, U. and Welzl, E. (2001). A continuous analogue of the upper bound theorem, *Discrete Computational Geometry*, Vol. 26, 205-219.

Wendel, J.G. (1962). A problem in geometric probability, *Math. Scand.*, Vol. 11, 109-111.

Appendix

Proof of Lemma 2: The numbers $A_{m,q,j}$ of (q,p)-paths such that $\max\{q - \underline{s}, p + \bar{s}\} < j$ remain to be computed. These computations are based on the 'reflection principle' due to Désiré André (see, e.g., Feller (1968), Chapter III). Let $A = (\alpha, a)$ and $B = (\beta, b)$ be points in the positive orthant: $\beta > \alpha \geq 0$, $a > 0$, $b > 0$. By reflection of A on the x-axis is meant the point $A' = (\alpha, -a)$ (see Figure 8).

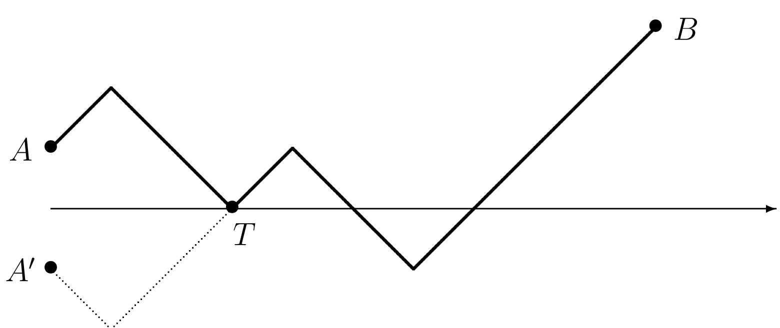

Figure 8

Reflection principle: The number of paths from A to B which touch or cross the x-axis equals the number of all paths from A' to B.

Let $N(m,c) = \binom{m}{\frac{m-c}{2}} = N(m,-c)$ denote the number of paths from $O = (0,0)$ to (m,c). Let a and b be positive, and $-b < c < a$. By the reflection principle, the number of paths from $(0,0)$ to (m,c) which touch or cross $y = a$ is equal to the number of paths from $(0,2a)$ (the reflection of O on the axis $y = a$) to (m,c), i.e., $N(m, 2a - c)$. By the same argument, the number from $(0,0)$ to (m,c) which touch or cross $y = -b$ is $N(m, c + 2b) = N(m, 2a - c - 2(a + b))$.

Now, by a double application of the reflection principle, a path from $(0,0)$ to (m,c) which touch or cross $y = a$ and then $y = -b$ (called an '(ab)' path in the sequel) can be first associated to a path from $(0,2a)$ to (m,c), itself associated to a path from $(0, -2a - 2b)$ to

(m, c); hence $N(m, c + 2(a + b))$ of 'ab' paths. A triple application allows through the same line of argument to denumerate the paths which touch or cross $y = a$, then $y = -b$, then $y = a$ again ('$(ab)a$' paths); their number is $N(m, 2a - c + 2(a + b))$. An extension of this method gives:

- $N(m, c + 2k(a + b))$ for the number of paths which touch or cross $y = a$ and then $y = -b$ k times in a row ('$k(ab)$' paths);
- $N(m, 2a - c + 2k(a + b))$ for the number of paths which touch or cross $y = a$ and then $y = -b$ k times in a row and then $y = a$ again ('$k(ab)a$' paths);
- $N(m, c - 2k(a + b))$ for the number of paths which touch or cross $y = -b$ and then $y = a$ k times in a row ('$k(ba)$' paths);
- $N(m, 2a - c - 2(k + 1)(a + b))$ for the number of paths which touch or cross $y = -b$ and then $y = a$ k times in a row and then $y = -b$ again ('$k(ba)b$' paths).

Our aim is to compute the number of paths from $(0, 0)$ to (m, c) which touch or cross neither $y = a$ nor $y = -b$. This comes first by exclusion of paths which touch or cross $y = a$ and paths which touch or cross $y = -b$. But thus '(ab)' and'(ba)' paths are excluded twice and must be reincluded once. But then '$(ab)a$' and'$(ba)b$' are excluded twice, then reincluded twice, and therefore must be re-excluded once... This standard application of the *inclusion-exclusion principle* (see Comtet (1974), Chapter IV), leads to the formula:

$$\begin{aligned} N(m,c) &- N(m, 2a-c) + \sum_{k>0}[N(m, c+2k(a+b)) - N(m, 2a-c+2k(a+b))] \\ &- \sum_{k>0}[N(m, c-2k(a+b)) - N(m, 2a-c-2k(a+b))] \end{aligned}$$

for the concerned number, which can be rewritten:

$$\sum_{k}[N(m, c+2k(a+b)) - N(m, 2a-c+2k(a+b))]$$

(over all integers k from $-\infty$ to $+\infty$, but only finitely many non-zero terms). The formula 1 obtains readily by substitution of the right parameters. ■